

The great sky greenhouse

Last week's hoop-la notwithstanding, the most urgent need in respect of global warming is to negotiate an international convention on greenhouse gases. *Nature* will play its part with regular summaries of relevant research.

LAST week was a famous week for the greenhouse effect. Mrs Margaret Thatcher, the British prime minister, effectively upstaged Working Group I of the Intergovernmental Panel on Climatic Change (IPCC) of the United Nations by announcing, six hours ahead of time, that Britain is so seized of the reality of the excess greenhouse effect — that proportion of near-surface temperature attributable to anthropogenic rather than natural causes — that it will arrange to restore the emission of greenhouse gases in 2005 to those in 1990 “provided that” this can be done within the framework of an international agreement. Such is Mrs Thatcher's reputation that she was not required to explain how this would be done.

Meanwhile, there are several important issues to keep clearly in mind, not the least of which is that even if the excess greenhouse effect is not yet real, it must at some stage be a factor to contend with. Pangloss, if he were with us still, would surely agree. For the past two years, this journal has been advocating that even the prospect that industrial and agricultural greenhouse gases may change the climate requires that urgent attention should be given to the negotiation of an international convention whose effect would be to limit them. That — an agreed convention with the numbers left blank, to be filled in at a later stage — remains the urgent need. What is being done?

IPCC has three working groups, concerned with the estimation of the extent and nature of likely climatic change (which reported last week — see page 373), with the consequences thereof and the means by which they might be abated. The full IPCC will meet in August. Its report will go to an intergovernmental two-week meeting beginning at the end of October, at which the draft of a convention on greenhouse gases will be drawn to every government's attention. The most important item of that agenda will be ‘Date of Next Meeting’. The best hope is that it will not be long delayed, for even a second meeting of the kind proposed cannot be definitive.

Meanwhile, there are several cautions to keep in mind. First, the magnitude of the predicted increase of average global temperature over the next half century is uncertain by a factor of three or thereabouts. (That is why IPCC's Working Group I was seriously in error last week in releasing a general summary of its findings to the general press before telling the professional community how it arrived at its conclusions.) Second, the group concerned with the consequences of climatic change has allowed

conditional to be replaced with indicative phrases: for “will be” read “might be”. Finally, it is plain that no working politician has more than the sketchiest idea what might be done to ameliorate the effects of an excess greenhouse effect, or to avoid them.

What the years ahead will bring is anybody's guess. Much will depend on the degree to which the computer predictions are borne out by measurements of climate on the surface and of the Earth's radiation balance from Earth satellites. But the most important area of relevant research is also the most neglected. The cost of either adaptation to or avoidance of the excess greenhouse effect may well amount to a substantial fraction of the world's total stock of social capital. Advocates of windmills or of nuclear power as a source of electricity will be much heard of in the years ahead, but the chief burden will fall on the economists who will have to say what are the costs of adaptation or avoidance.

Research, in what Anglo-Saxons call social as well as natural science, will nevertheless play a crucial part in the assessment of what must be an unfolding set of circumstances. This journal is glad to have published in the past few years many of the key documents in this field, but there are many other, perhaps more specialised research reports that naturally find their way into other journals. Beginning next week, *Nature* will therefore publish occasional articles and information designed so as to help readers judge for themselves the quality of the research being carried through in this interesting and important field, its import and the use being made of it by policy-makers. Authors of original research, or of research summaries intended for use by committees, official and otherwise, are invited to send copies to the Editor of *Nature*, indicating clearly when their contents may be made public without embarrassment. □

German questions

Reunifying the two halves of Germany, could come unstuck, with serious consequences for the rest of us.

THIS week's summit in Washington will make it plain that Mr Mikhail Gorbachev, the President of the Soviet Union, has been compelled by events, by the reservations of his opponents and by the diffidence of his friends to end



his compliance with the project to merge the German Democratic Republic (DDR) with the Federal German Republic (FRG). Gorbachev has a point: while arrangements for the security of Central Europe are in the air (and while his own position in the Soviet Union is unclear), the reunification of Germany is premature, to say the least.

Even so, the reunification of Germany is also inevitable. East Germans have for a year been able to move West if they chose. Up to 2,000 a week have been taking advantage of the opportunity. At that rate, the DDR would be emptied of people — except for the very old and, possibly, the very young — in about a decade. The result of that would be a social and military vacuum in Central Europe, especially in the eastern half. West Germany's Chancellor, Helmut Kohl, campaigned strongly during last December's East German elections, but he was merely acting prudently, seeking to avoid an intolerable dilemma.

Yet the pace of events is too fast for comfort. Those who have been engaged in relatively much simpler projects than the merger of whole countries — the merging of, say, departments of physical and inorganic chemistry — know that whole years may go by before a consensus is established. How can it be possible to merge two whole countries in a few months, which is the present intention? Only by taking a whole raft of considerations on trust, by agreeing in advance to agree on certain principles. The trouble with German reunification is that the only tangible agreement so far struck is that up to 6,000 marks of individual personal savings by East Germans will be converted into Deutschmarks at par, with the excess converted at the rate of two for one. It is also agreed that Ostmark salaries will be converted at face value into Deutschmarks, but that is a fiction: either the enterprises at which those concerned work are able to make their way in the competitive world, or they cannot. In the second case, the employees will have to look for work in West Germany, which is the reason why the German question arose in the first place.

The confusion is especially apparent in research. The West German grant-making agency called the *Deutscher Forschungsgemeinschaft* has said that it will, from the beginning of next year, allow East German applications for funds to be considered alongside those from the West. But there is deep disagreement in West Germany about the criteria that should apply. Should the same standards apply to both, or should East German applicants be given an inside track? The first case would spell the end of research in East Germany, the second would breach the principle of excellence on which West German research has prospered in the past decade. Nobody knows how that will come out, nor what the *Max-Planck Gesellschaft* is up to. What it should be doing, of course, is to accept that the East German Academy of Sciences should be a titularly equal partner until things can be sorted out. In general, that is what the circumstances demand — time to sort things out. □

Labour's ambiguity

THE British Labour Party, which hopes to be the next government, should have a policy on research.

THE British Labour Party, temporarily known as Her Majesty's Opposition (but which should constitutionally prefer to be known as Her Majesty's Government's opposition, where the capitalization of nouns is intended), last week produced its provisional agenda for the next British general election. The document is a frank disappointment. On higher education and research, there is only one tangible proposal — the funding councils that at present provide public money to universities and polytechnics will be merged (a sensible idea, on the cards since the late Mr Anthony Crosland invented British polytechnics in 1968). Otherwise, the document says nothing but that there is a need for skill and, therefore, for training as well as for equity in access thereto.

Even calculations of electoral advantage would have commanded a different tack. The present government, in office for 11 years, has ruined the capacity of British research to be imaginative by requiring it (in the proper public interest) to count pennies instead. British academics as different as aristotelians and theoretical physicists have become adept at the allocation of car parking spaces and, by all accounts, exceptionally skilled at telling when the government's apparently perpetual offer of early retirement on favourable terms will allow them to combine a half-time appointment at an American university with growing roses in a clement English county. Demographically, of course, it is even more troubling that younger scholars increasingly regard their first overseas postdoctoral appointment as their chance to play in a different league. The Labour Party's provisional manifesto says nothing on this point, and little to suggest that it would try to make Britain a more interesting place for researchers, in that role, to work. This is all a far cry from the then Mr Harold Wilson's "white-hot technological revolution". There were serious faults in that programme, but the cry was heartening, and electorally appealing.

Last week's document is similarly thin on research as such, which comes through as a handmaiden of industrial policy. The Labour Party (like the present government) would cut back on defence research. But the party's promise, at the last election, that research would be paid for by an industrial levy, is mercifully not repeated. It is nevertheless crucial to those working in research that there should be some undertaking of the scale and the direction in which research funds would be channelled. What the Labour Party should do, between now and the time that electoral considerations will require that it should publish a more substantial successor to last week's document, is to decide what science is for. The party evidently believes, on the evidence of last week's document, that research is either for wealth-creation (the present government's belief) or job-creation. Is that all? □

Modest response to climate change threat

■ Little new action proposed

■ IPCC process questioned

London

AN early draft of the Intergovernmental Panel on Climate Change (IPCC) report on responses to global warming which was leaked to the press last week has outraged environmentalists by failing to call for stringent cuts in greenhouse gas emissions. But the outcry may persuade IPCC's working group III to take a tougher line and revise the report at its final meeting in Geneva next week.

On the key issue of carbon dioxide emissions, the draft report does not suggest major changes from current national positions. Western European countries, the report says, may be able to "stabilize or reduce carbon dioxide emissions by early in the next decade". No mention of reductions is made for other regions. Eastern Europe and the Soviet Union may, at best, stabilize emissions over the next two decades, and developing countries may be able to "reduce the annual growth in carbon dioxide emissions from over 3 to around 2 per cent", the draft says.

It is the suggested response from 'North American and Pacific OECD countries' that has attracted most criticism. The report says only that the United States and Japan may be able to slow the growth in carbon dioxide emissions. No timetable is given.

Simon Roberts from Friends of the Earth says "this looks like political intervention with no science at all", a view shared by most environmental pressure groups. Unlike IPCC's other two working groups, which consist mostly of scientists, group III is chaired by a US representative, dominated by civil servants, and has strong Japanese representation.

Group III's suggestions, if carried out, would slow the growth of global carbon dioxide emissions. But the atmospheric concentration of the gas would continue to rise. Climate models produced by IPCC group I, whose report was finalized last week at a meeting in the United Kingdom, predict that emissions of long-lived greenhouse gases (including carbon dioxide) must be cut by more than 60 per cent to stabilize atmospheric concentrations at today's levels.

The group I report predicts the climate change that may result under several different scenarios. If carbon dioxide emissions increase unchecked and atmospheric carbon dioxide concentration doubles by 2025–50, global mean temperature is predicted to increase at about

0.3° C per decade. The most rigid anti-greenhouse strategy examined, reducing carbon dioxide emissions substantially over the next century, gives a predicted rate of increase of about 0.1° C per decade. The implication is that some degree of economic adaptation to climate change, regardless of the action taken to reduce the greenhouse effect, is now probably inevitable.

An early draft of group II's report, looking at the impacts of climate change, was also leaked to the press last week. The draft, to be finalized in Moscow this week, assumes no reduction in carbon dioxide emissions and makes several alarming predictions. Millions of people may be uprooted by flooding, erosion and changes in agriculture, into areas that are "likely to have insufficient health and other support services". Group II's draft says that global agricultural production could be sustained if there are no reductions in greenhouse gas emissions, but many natural ecosystems could not cope with the expected rate of climate change.

Ralph Cicerone, from the University of California, Irvine, one of the scientists who worked on group I's report, believes expectations from "a hasty study like CARBON DIOXIDE EMISSIONS

IPCC" may have been too high. He contrasts the IPCC process, which has been running for little over a year, with the long haul towards the three-volume report on chlorofluorocarbons and the ozone layer, which formed the scientific background to the Montreal protocol. The US National Aeronautics and Space Administration began work in 1977, but the final report, compiled with the World Meteorological Organization, was not published until 1986. Cicerone thinks that IPCC must be regarded as a first step towards international agreement, not the final answer. But he hopes that politicians will not use the IPCC reports as "an excuse to do nothing".

A lack of communication between IPCC's working groups may have caused problems. One scientist involved with working group I suggests that the whole process would have been better conducted in serial, rather than parallel with group I feeding its completed report into group II, and so on.

Alan Miller, from the Center for Climate Change at the University of Maryland, thinks that the attitude of the British prime minister, Mrs Margaret Thatcher will now be a key factor (see below). West Germany, the Netherlands and Scandinavian countries are pushing for cuts in emissions, but Britain has been seen to support the United States. IPCC will present a single report to the World Climate Conference in the Autumn, so there will further scope for amendment of group III's position on greenhouse emission controls.

Peter Aldhous

British pledge on reductions

Bracknell, Berkshire

PRIME minister Mrs Margaret Thatcher last Friday (25 May) pledged Britain to return its carbon dioxide emissions back to 1990 levels by 2005, provided the target was part of "a wide international effort". This involved undercutting present UK emission projections by "up to 30 per cent", she said.

Mrs Thatcher was in Bracknell to open the Meteorological Office's new Hadley Centre for climate prediction and research. Dressed in a dark suit matching her sombre message, she spoke of the flooding and drought that could result if global warming is not checked. "People will be crying out, not for oil wells, but for water", she warned.

But the speech gave little detail of how British emissions would be stabilized. An overhaul of transport and energy policy is thought essential by many, but Mrs Thatcher chose instead to highlight the role of the private sector. Worldwide application of a new method of ammonia production, developed in Britain, would have the

same effect on carbon dioxide emissions as removing five million cars from the road, she said. Details of government policy on global warming will be given in a white paper (policy document) on the environment, in the autumn.

Dan Lasher, from the US Natural Resources Defense Council, countered immediately after the speech that the prime minister's proposal on UK emissions "doesn't respond to the threat she underscored". If developed countries merely freeze emissions, he said, global emissions will still increase with growing energy demand in developing countries.

The proposed British response fits within the loose guidelines on emissions laid down in the leaked draft report from working group III of the Intergovernmental Panel on Climate Change (see above). But other Western European countries plan reductions — West Germany has already committed itself to a 25 per cent reduction in emissions from 1987 levels by 2005.

Peter Aldhous

A solution in iron solution?

Washington

A proposal that the greenhouse effect could be inexpensively mitigated by artificially stimulating massive blooms of algae in the Antarctic Ocean has outraged environmentalists and drawn attention from researchers, the press and even the White House. The proposal, by John Martin of the Moss Landing Marine Laboratory was first published in these pages (see *Nature* 345, 114 & 156; 10 May 1990), but when the *Washington Post*, in a front-page story, independently picked up the idea last week, Martin's office was besieged by calls.

"I'm astounded", Martin says of the reaction. However, the National Research Council took his idea seriously at a workshop convened last December at the suggestion of Adam Heller of the University of Texas. During the two-day meeting, participants even suggested a \$50 to \$150 million trial experiment.

The logic of Martin's idea comes from evidence that the growth of phytoplankton in the Antarctic Ocean is restricted by a shortage of iron. Other nutrients abound, but in the southern regions far from dry land and wind-borne dust, Martin finds that levels of dissolved iron are among the lowest in the world. Because phytoplankton fix carbon dioxide during photosynthesis, adding about 300,000 tonnes of iron to the ocean could stimulate algal growth sufficiently to soak up more than 2,000 million tonnes of carbon — roughly a third of the amount put in the atmosphere each year by deforestation and fossil-fuel burning.

For those who worry that enforcing drastic cuts in CO₂ emissions would return the developed world to the pre-industrial age, man-made algal blooms sound like an attractive way out. But Martin's solution to climate change is attracting plenty of criticism.

Biologists, noting the deadly effects of algal blooms in the Adriatic, are concerned about the impact on the food chain. Oceanographers and atmospheric scientists say that the phytoplankton may simply release most of the carbon back into the atmosphere when they die, rather than sinking to the ocean floor. And some policy analysts worry that if every industry could fill a tanker with iron to atone for its carbon sins, few would take expensive and restrictive measures to keep emissions down. "We're not iron-poor, we're pollution-rich", says Daniel Becker of the Sierra Club.

Environmentalists look poorly on 'megabiology' — global tampering with the biosphere. Atmospheric scientist Michael Oppenheimer of the Environmental Defense Fund is reminded of earlier suggestions to combat global

warming with huge mirrors in space (to reflect away sunlight) or to regenerate stratospheric ozone with orbiting lasers. "The whole thing smacks of the mega-project-itis that infects large areas of the scientific community when it comes to environmental problems", he says. "If there is one thing we've learned in the last twenty years, it's that there are always unintended implications of what we do to the environment."

Reaction from the scientific community has been more encouraging, although serious questions remain unanswered. Among the more glaring gaps are the limited understanding of CO₂ exchange between the atmosphere and the ocean and the movement of carbon from the surface to greater depths. Although Martin says iron appears to be by far the most important limiting agent for algal growth in polar water, his is the only major study on the subject. When saturated with iron, large-scale phytoplankton blooms may still be limited by some other unexpected nutrient or condition. "Biological systems are famous for not doing what they're supposed to do", says David Caron of the Woods Hole Oceanographic Institution.

Others suggest that Martin may be observing an increase in the rate of algal growth, rather than its overall volume.

FRENCH UNIVERSITY REFORMS

Emergency investment plan under way

Paris

A CABINET meeting on 23 May approved French education minister Lionel Jospin's plans to invest an extra FF16,000 million (\$2,900 million) over five years in order to relieve severe and growing overcrowding in France's universities (see *Nature* 342, 723; 1989).

The move follows a promise by President François Mitterrand earlier this month to build 1.5 million square metres of new university facilities. Mitterrand said last week that the new plans are "unprecedented", but added that so is the demand. Student applications for university places were up by about 80,000 last year and the extra demand is expected to be even greater this autumn. Early last year, Jospin declared that 80 per cent of pupils should leave school with a university entrance certificate (*baccalauréat*).

University buildings in the Paris region are already severely overcrowded and several are showing signs that they were built in a hurry to meet a similar expansion in the early 1970s. Five new universities are planned for suburbs and dormitory towns around Paris, another for Pas-de-Calais on the Channel coast and possibly

But Martin says laboratory tests indicate that the addition of iron can increase total phytoplankton biomass tenfold. And even if it is only the rate of growth that is accelerated, short Antarctic summers put a premium on speedy development. During the few months when polar conditions are right, the fastest-growing algae are likely to produce the most biomass and soak up the most CO₂, he says.

Despite the doubts, the scientific consensus appears in favour of trying out Martin's theory in the field, on a modest scale. Martin suggests putting a tonne of iron in an area 10–100 km square in the Gulf of Alaska, which is relatively easy to get to and has phytoplankton conditions similar to the Antarctic. An experiment of that sort would cost between \$10 and \$50 million, he estimates.

Such a test could answer the most immediate questions about the impact on other plants and animals in the region of the algal blooms. The NRC is expected to convene another workshop on the subject in August or September. NRC officials say a full-scale, peer-reviewed report could be released in late 1991. Along with the National Science Foundation and the Electric Power Research Institute, which funded part of the NRC workshop, the US Environmental Protection Agency has expressed interest in funding further research in the area as part of the \$1,000 million US climate change programme.

G. Christopher Anderson

two others in the provinces.

There is also a critical need to extend existing facilities, especially libraries and student accommodation. But the government's new package is unlikely to be sufficient to cover all the costs. So Jospin is relying upon regional and local governments to share some of the burden. This is may not be easy, especially since the government is unwilling to let local authorities have much of a say in higher education policy.

The plans have received a lukewarm reception from university teachers' unions. The new funds are felt undeniably to be welcome but are inadequate, they say. Also they fear that local governments will not pick up the extra bill because they will be reluctant to run the electoral risk of having to pass it on to the taxpayer.

Jospin's plans are still not out of the woods. The education budget has to be approved by parliament in the autumn, although the proposals have received guarded opposition support. A series of hearings are also planned this summer to negotiate the university reforms in each region and administrative department of the country.

Peter Coles

High hopes for Astro-1

Washington

IN 1978, when the space shuttle was being advertised as a regular bus service into Earth orbit, the National Aeronautics and Space Administration (NASA) invited scientists to propose instrumental payloads that could be flown repeatedly, once or twice a year, making the shuttle into an easily used space science platform. But it has taken 12 years for the first of those proposals, a set of astronomical telescopes known as Astro-1, to get off the ground. And the future of Astro, due to have been launched yesterday, 30 May, is far from guaranteed. No additional missions are scheduled, and only after this first flight has been evaluated will NASA decide whether the instruments, tailor-made for the shuttle, will be used again.

Astro-1 consists of two separate astronomical packages which find themselves on the same flight because of the disruption caused by the 1986 Challenger explosion. The larger part of Astro-1 is a set of three ultraviolet telescopes — an imager, a spectrophotometer and a polarimeter — that was to have flown in early 1986, along with a camera specially designed to photograph Comet Halley. When shuttle flights resumed three years later, Halley was long gone, and instead of the camera the Broad-Band X-ray Telescope (BBXRT) was mounted alongside. BBXRT was originally part of an X-ray astronomy package, but was pulled out and made ready earlier in order to take X-ray pictures of supernova 1987A.

The two parts of the Astro-1 package are on separately steerable mounts attached to pallets inside the cargo bay. As the shuttle pilot holds the craft for several hours in a fixed attitude, other crew members point the telescopes at a series of predetermined target stars. All three ultraviolet instruments — the Hopkins Ultraviolet Telescope (HUT), the Ultraviolet Imaging Telescope (UIT) and the Wisconsin Ultraviolet Photo-Polarimeter Experiment (WUPPE) — are aligned together and observe the same object, and for about 20 per cent of the planned observations BBXRT will also be pointed in the same direction to provide simultaneous X-ray imaging. Both ultraviolet and X-ray astronomy specialize in energetic cosmic phenomena: young stars, supernova remnants, quasars and hot interstellar gas.

With the exception of UIT, which takes photographs that will be developed on the ground, the Astro-1 instruments will send down data in real-time, allowing astronomers at Marshall Space Flight Center in Alabama to re-adjust the observing program should some transient phenomenon such as an X-ray binary or a cataclysmic variable erupt. According to Peter

Serlemitsos, BBXRT project investigator, the enforced hiatus caused by the Challenger accident has allowed mission controllers to become much more comfortable with last-minute schedule changes.

In the weeks following the ten-day mission, astronomers will be more than usually anxious to see what their search of the sky has trawled up. Although NASA has promised money to support maintenance and even some further development of the Astro-1 hardware, according to HUT co-investigator Knox Long, no commitment even in principle to another launch has been made. Both Long and

SUPERCONDUCTING SUPER COLLIDER

Pot of gold sought in the East

Washington

ARMED with little more than a letter from President Bush and a blueprint, a US Department of Energy (DOE) team departs for Japan and Korea this week to find partners for the Superconducting Super Collider (SSC).

Although DOE officials do not expect to come back with any contracts, they do hope to start work on government-level collaboration agreements. However, what DOE is willing to offer in exchange for money or equipment remains a 'classified' secret. Some Japanese researchers have expressed interest in the project, but the government of Japan has sent few public signals of encouragement.

DOE officials have previously indicated that the international foray would begin as soon as a final cost estimate for the 20 TeV proton accelerator had been set. But as DOE deputy secretary W. Henson Moore and other DOE and White House officials prepared to leave last week, that price was still only a guess of around \$8,000 million.

Congress recently passed legislation demanding that one-third of the project's cost come from non-federal sources. The state of Texas has promised \$1,000 million, leaving DOE still to find an additional \$1,600 million or so elsewhere. So far, the only definite international promise had been \$50 million from India, but even that appears to have gone soft. "We are not sure how solid any of those past commitments are. We're treating them all as tentative", Moore said.

Whether DOE will allow foreign partners to make the more sophisticated components of the accelerator — to the advantage of their high-tech industries — remains an open question. Congress has stressed that technological manufacturing, particularly of the SSC's superconducting magnets, should remain at home (for some politicians, technology

Serlemitsos are expecting NASA to evaluate Astro-1 later this year, with the payloads being kept in storage at the Kennedy Space Center until, as Long puts it, "the situation clarifies". Another mission could be put together at about six months' notice, provided everything works as planned, but the nearest hint of an open slot in the shuttle manifest, says Serlemitsos, is not until 1993.

The possibility of avoiding these problems by making the two parts of the Astro mission into self-contained satellite observatories, needing the shuttle neither for launch nor for operation, is now almost gone. Astro-1 may have to do spectacularly well to avoid premature consignment to a museum.

David Lindley

spin-offs remain the project's chief selling point). Moore announced that the contract for the magnets is now up for grabs, and both US and international companies may apply. But whether Japan will be given a real chance at the contract as part of a collaboration deal is not known.

DOE officials plan to extend their search for partners to Europe in October. By then they also hope to have named a director for the project at DOE headquarters and to have settled on a final price.

G. Christopher Anderson

Tokyo

Yubun Narita, director of scientific affairs at the Japanese Ministry of Foreign Affairs, says that the meeting will be the first detailed explanation of the project by US government officials, and as such Japan cannot make any commitments at this stage. Only after hearing the US explanations and requests for help will Japan decide what (if anything) to do next.

The views of Japanese high-energy physicists on the SSC are divided. Some researchers at the High Energy Physics Laboratory (KEK) in Tsukuba are keen to participate in the experimental phase of the programme. But others in the Universities would prefer to collaborate with the European Laboratory for Particle Physics (CERN) in the operation of CERN's proposed Large Hadron Collider because participation in that (if it goes ahead) is expected to be cheaper.

At present, there seems to be little support among the Japanese scientific community for Japan to contribute to the construction costs of the SSC. One physicist at KEK asks "What is the point if they don't want our superconducting magnets? Would our contribution be some form of ODA (Overseas Development Assistance)?"

David Swinbanks

Protesters ask for more

Washington

Two hundred police officers wearing riot helmets and rubber gloves were called in to protect the US National Institutes of Health (NIH) last week as a crowd of nearly 1,000 activists massed on its campus demanding an acceleration in AIDS research, increased access to experimental drugs for minority groups and more vigorous research into AIDS-related opportunistic infections.

About 60 demonstrators were arrested on the campus for attempting to break into buildings and a further 21 were arrested while staging a sit-in in the office of Daniel Hoth, director of the division of AIDS for the National Institute of Allergy and Infectious Diseases (NIAID). But despite fervid chants of "We want a cure, this is a war" and "storm the NIH", there were no reports of serious injuries or damage to property, with most demonstrators being content to wave placards or adorn NIH trees with symbolic red tape.

Anthony Fauci, director of NIAID, described the protest as "street theatre" — interesting, but unhelpful. "It was rather inappropriate that they should demonstrate against the very institution that is trying to help them", he said.

Organized by the militant US AIDS activist organization ACT-UP, the demonstration marks the latest battle in an increasingly emotional campaign by activists to influence the course of AIDS research. Mark Harrington, ACT-UP's spokesman on research policy, claims that the demonstration was precipitated by a growing feeling among activists — that NIH policies on drug development are "too conservative and unimaginative", and that there are "too many drugs coming out of the test-tube but not getting into bodies". The last straw, says Harrington, was hearing that the initiation of new clinical trials of drugs would be delayed by nine months because the database was switched from the Research Triangle Institute at North Carolina to the Statistical and Data Analysis Center at Harvard Medical School.

But Fauci dismisses that claim as one of a number of "absolutely untrue" allegations — "the same rate and pace and number of trial protocols have been put into effect" this year as were executed last year, he says.

NIAID's policies on research and clinical trials have been bitterly attacked in documents released both independently by ACT-UP and jointly by 25 different AIDS activist organizations, ACT-UP among them. One of the key criticisms is that NIAID's clinical trials, which at present enrol only 1 per cent of US AIDS patients, are too limited and exclude minority groups such as children and

pregnant women. Fauci, however, says that the activists fail to acknowledge that the need for increased medical services for minority groups in the United States is part of a "general problem in the medical-care delivery system". He argues that NIAID has already introduced a host of measures designed to help increase the accessibility of clinical trials to minority groups and points out that 15 of its trial units are now devoted to paediatrics.

The documents also claim that NIH-directed research into the many opportunistic infections that afflict AIDS patients is inadequate, particularly as some AIDS patients are now living longer because of AZT. Fauci says that NIAID is "clearly aware of the fact that opportunistic infection research has not proceeded as quickly and as intensively as research on HIV anti-retroviral drugs" and has already taken steps to "correct that" by expanding its trials of treatments for opportunistic infections. The number of patients involved in such trials has more than tripled in the past year, he says.

EC SCIENCE

Worries over Framework

London & Munich

THE emergence of the European Communities' (EC) third Framework research and development programme last month, with a shift towards more basic science, has provoked apprehension in the UK research councils that British research will suffer as resources switch to EC projects. The new Framework Programme is set to run from 1990 to 1994 and has a budget of 5,700 million ECU (1 ECU = £0.72), an increase of over 300 million ECU over the previous five-year programme. But very little of the new programme's budget will be spent until 1991, and the previous programme runs into 1992.

In Britain, the Treasury adjusts the budgets of spending ministries to account for the United Kingdom's contributions to (and returns from) the EC. This 'attribution' depends on secret negotiations before the annual autumn statement on public expenditure. Similar arrangements exist in other member states, including West Germany, but they are less rigid. "Our contributions to the EC have been growing for the past 10 years and BMFT [research ministry] budgets have never been cut", says a BMFT spokesman.

'Attribution' worries the UK research councils because as the new Framework programme increases support for basic research, the Department of Education and Science (DES)'s budget, from which research councils are funded, may be squeezed. A total of 518 million ECU is

One down, one to go

London

THE UK agriculture ministry has announced that a 15-month-old cow suspected of having bovine spongiform encephalopathy (see *Nature* 345, 280; 24 May 1990) did not have the disease. But the ministry says that its Central Veterinary Laboratory at Weybridge is still examining a second suspected case in a cow born after the ban on cattle feed containing ruminant tissues.

Peter Aldhous

The other main allegation is that in setting their research priorities some NIH researchers are being swayed by the need to please the drug companies that top up their coffers. The activist documents demand that principal AIDS researchers be bound to disclose their financial ties just as members of FDA advisory committees are. But Fauci says the demand is unnecessary — AIDS researchers must already divulge their financial interests, and the staff of NIAID, who ultimately decide which drugs should be tested, are forbidden to have any ties to industry.

David Concar

planned for basic science under the new Framework programme, increased from 180 million ECU spent from 1988–92.

A Science and Engineering Research Council (SERC) memorandum to the House of Commons Education, Science and Arts Committee says that it is difficult for SERC to assess EC research programmes "without knowing whether resources . . . would come from the [DES] science budget". Dai Rees, secretary of the Medical Research Council, fears that the DES 'flexibility margin' — a fund of about 15 million ECU kept back for priority research arising after the rest of the science budget is allocated — may be swallowed up by attribution against EC research spending. Rees is critical of EC medical science, also set to grow under the new Framework programme. A tendency to "throw money at things to look good" exists in the EC, he says, "which goes against the grain when you're turning down good projects".

Sir Peter Swinnerton-Dyer, chairman of CODEST, the committee that oversees EC support for basic science, says Rees's concern over attribution against the DES budget is valid "in principle". But he thinks Rees overestimates the problem. Eileen Buttle, secretary of the Natural Environment Research Council, says the research councils must ensure that their views are heard when EC research programmes are drawn up.

Peter Aldhous & Steven Dickman

Superpigs go to market

Sydney

THE Australian government is under pressure to introduce new laws to govern the regulation of genetic engineering after a public outcry over the sale for human consumption of 53 genetically altered pigs. The pigs had come from a research programme aimed at breeding 'superpigs' through the introduction of extra copies of the gene for porcine growth hormone. They were sold for slaughter apparently without approval from any of the bodies regulating genetic engineering.

At present, guidelines on genetic engineering in Australia are issued by the federal government's Genetic Manipulation Advisory Committee (GMAC). Any organization seeking to release a genetically altered organism into the environment is expected to seek advice from GMAC, although compliance with the guidelines is voluntary.

Metrotec Pty Ltd was responsible for the experiments with the pigs. The 53 pigs they sent to the slaughterhouse were the offspring of a transgenic boar not expressing the gene and had been removed from the programme because their growth hormone gene was not functional. The company is part owned by the University of Adelaide, but the university's Institutional Biosafety Committee had not assessed the dangers of releasing the pigs.

Barry Lloyd, managing director of Metrotec, said on ABC Radio that while the university is required to conform to the guidelines on genetic engineering, private companies are not.

Metrotec further argues that the pigs were released with the knowledge of both state and federal government bodies. John Smeaton, a director of Metrotec, says that the National Health and Medical Research Council and the South Australian Health Commission had approved the meat for consumption. Smeaton also says that the release of the pigs had been

discussed with Brian Booth, a representative of the Recombinant DNA Monitoring Committee, the precursor to GMAC.

According to Lorne Skene, project manager of the genetic manipulation section of the Victorian Law Reform Commission, GMAC has no powers of direct sanction should their guidelines be ignored. GMAC can, however, recommend the removal of federal funding and tax incentives from both the research group involved and the parent institution.

The Australian Conservation Foundation (ACF) is taking advantage of the questions raised by release of the pigs to push for a national uniform system of legal regulation. Phillip Toyne, director of ACF, said: "As companies increasingly seek to commercialize research results new laws are becoming even more necessary." The University of Adelaide also believes GMAC guidelines should become enshrined in legislation. "This would ensure that all organizations operating in this area would be required to meet defined standards", the university registrar said in a public statement.

In a report released last year, the Victorian Law Reform Commission recommended legislation requiring mandatory notification of any proposed release of a genetically altered organism to GMAC and to any relevant state or federal body, such as the Department of Agriculture.

The commission also called for an environmental assessment before the release of any experimental organisms and the advertising of proposed releases "to ensure that interested people can obtain information and participate in decision-making before the proposal is approved". Australia's 'superpigs' were the world's first transgenic animals with enhanced production characteristics. In a project run by Bob Seamark, reader in endocrinology at the University of Adelaide, transgenic pigs incorporating genes for porcine growth hormone grew as much as 17 per cent faster, with a 30 per cent better feed conversion rate.

Seamark says he is against the introduction of legislation, "because it is inflexible and irreversible, [and] could close down the whole of recombinant research in one act of parliament. The solution is to develop confidence in researchers who abide by voluntary guidelines. They are not a bunch of naive wackers doing strange and evil things." Metrotec, however, would welcome legislation. "We would be happy with a stable set of rules to which we would have some input. We don't want to invest five years work into these animals only to have someone reinterpret the voluntary guidelines and say we can't market the pigs", Smeaton says.

Tania Ewing

Gorbachev leads silicon summit

Washington

SOVIET President Mikhail Gorbachev and other Soviet officials will visit Stanford University and research centres in Silicon Valley in California next week to lay the groundwork for a similar science and high-technology oasis near Moscow. The visit will begin a ten-day meeting between the two countries to establish trade relations, with the aim of encouraging US companies to participate in a 'Soviet Silicon Valley' in the city of Troitsk, near Moscow.

Known as the Soviet Silicon Summit, the meeting will follow the Bush-Gorbachev summit at which the United States is expected to grant the Soviet Union 'most favored nation' status. US officials have also declared their intention to lift many of the restrictions imposed by the Coordinating Committee on Multilateral Export Controls (COCOM) on the sale of high-technology products to the Soviet Union.

Although Gorbachev is expected to spend only one day at Stanford, visiting the university and meeting Silicon Valley business leaders, other Soviet officials will stay to tour the area's laboratories and production facilities. Next week, the officials are expected to unveil plans to convert Soviet military technologies to commercial use in fields such as microelectronics and communications.

As part of their drive to beat swords into microchips, Soviet officials have recently signed an agreement with Bechtel International, a large US engineering company, to study the feasibility of developing a 'technopolis' near Troitsk, which is already home to the Kurchatov Institute of Atomic Energy. The Troitsk Research Development Project aims to attract international businesses to help develop and invest in Soviet advances in fields such as nuclear fusion and laser technology.

The Soviet Union is expected to spend about \$100 million a year to support the project. If plans are approved, Bechtel will eventually build conference centres, a new international airport and an advanced telecommunications network in Troitsk by the year 2020. By that time, Troitsk, which at present has a population of 30,000, is expected to have a population of over 100,000.

During the Silicon Valley summit, to be held from 4 to 14 June in Santa Clara, California, Soviets will meet US companies interested in a range of business connections. The 15-member Soviet delegation is the largest to visit the United States to forge ties with US high-technology companies. After the meeting, the Soviets will open a full-time Silicon Valley sales office featuring demonstrations of Soviet software, Soviet market research data and a research library.

G. Christopher Anderson

Correction: Dounreay

THOMAS Robertson of the Hahn Meitner Institute (HMI) was misquoted in the article "Waste issue threatens reactor" on 10 May (*Nature* 345, 105; 1990) as saying that fuel from a West Berlin research reactor is not kept separate from nuclear bomb material at the Dounreay facility in Britain; HMI has since been assured that spent fuel from HMI would not be reused for military purposes at Dounreay. □

Correction: In vitro fertilization:

AN editing error in last week's news item on an MRC study on IVF (*Nature* 345, 283; 1990) changed the meaning of the last sentence, which should read: "In the opinion of Brian Lieberman, from St Mary's Hospital in Manchester, improved preservation of frozen embryos over the next two or three years should mean that multiple births will be less common. . . ." □

Success coming — but only slowly

Tsukuba

A RUSH by private companies to build research institutes is finally bringing Tsukuba science city to life after nearly two decades when it seemed Tsukuba might never amount to anything more than handful of government research institutes spread out over a 20 km strip of paddy fields.

New statistics from a Tsukuba city office show that since 1985 the number of private research laboratories has more than tripled from fewer than 30 to more than 90. Growth has been particularly dramatic in the past year: 18 new institutes opened in 1989, many more are under construction or planned and the number of company researchers and their support staff has risen dramatically to 5,000 — close to the number (6,700) of government researchers.

The newcomers include giant pharmaceutical and chemical companies, such as Yamanouchi Pharmaceutical, Sankyo, Takeda Pharmaceutical, Kyowa Hakko, and Sumitomo Chemical, and electronic manufacturers such as NEC. Some of the biggest laboratories belong to foreign companies (see opposite). And although much of the research is product-oriented, there has also been a blossoming of

company institutes devoted to much more basic research (see below). Around 50 small- to medium-sized factories producing high-technology products have also come in with the wave of private investment.

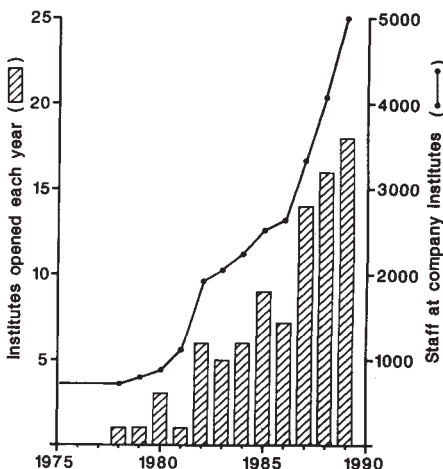
The surge of investment is good news for the government, which has spent more than a million million yen on a city in which until recently few people wanted to live. Plans to establish a science city in Tsukuba were hatched in the 1960s when many government research institutes located in Tokyo were rundown and overcrowded and expansion was impossible because of skyrocketing land prices. Hence came the idea to make a mass move to a new site outside Tokyo. The student riots of the late 1960s and early 1970s also accelerated moves to establish a new university in Tsukuba.

By the early 1980s, 46 government research institutes and two universities had moved to the new site. But they were spread out over a vast area. The 'city' lacked a shopping centre and cultural facilities. The nearest railway station was a 20-minute bus or expensive taxi ride away. And the local country schools could not meet the exacting standards of academic parents. The result was that many

researchers chose to spend four hours a day commuting from Tokyo rather than live in Tsukuba.

The situation is rapidly changing now. The influx of private researchers and their families helped to boost the population of Tsukuba city from 127,000 in 1985 to over 140,000 last year (the population is more than 174,000 if one includes outlying areas of the science city). New schools have been established to cater for the incomers. And although Tsukuba still has vast tracts of farm land, there is a shortage of land for housing — Japan's tax system encourages farmers to hold onto their farms.

Land prices have more than quadrupled in the past four years to over ¥1,000,000 (\$6,250) per tsubo (3.3 square metres) forcing some poorly paid researchers to set up house in country villages an hour's travel away. Many of the government researchers who previously scorned living



The histogram shows the number of private company research institutes opened each year in Tsukuba, and the graph shows the population of private company researchers and support staff in the city.

in Tsukuba are now looking for land and this has also helped drive up the prices.

So why the sudden rush to Tsukuba? The government-sponsored Tsukuba Science Expo (see *Nature* 314, 213–220; 1985) held in 1985 at a cost of ¥210,000 million (\$1,300 million at current exchange rate) acted as a great spur to private investment. Large department stores and hotels were built to cater for the 20 million visitors to EXPO, helping make Tsukuba a more liveable place. Also several of the companies have built their laboratories on the former EXPO site where cheap land was available.

Company officials cite two main reasons for their move to Tsukuba: availability of large tracts of cheap land not far from Tokyo, and access to facilities and information at government institutes. The synchrotron radiation source at the High Energy Physics Laboratory has many

Freeing the imagination

NEC'S new laboratory in Tsukuba provides an insight into the fundamental changes that are going on in the research and development policies of major Japanese companies. The laboratory is vast by Japanese standards, providing an average of 100 square metres of space for each of the 160 staff plus all of the latest equipment. But more surprising is that much of the research going on there is not of immediate use to a major electronics and computer manufacturer.

Researcher Shigeru Tanaka, for example, is trying to understand how the brains of cats and monkeys process visual information using a mathematical model that describes the domain structure of magnetic thin films. (Tanaka sees striking similarities between the patterns of up-and-down magnetization in thin films and the stripes of ocular dominance of the right and left eye in the visual cortex of the brain.) Toshikazu Takada spends his time simulating collisions between hydrogen molecules on a supercomputer. And another researcher plans to use pulsed lasers to investigate chemical reactions.

Such basic research is something completely new at a Japanese company laboratory. Hisatsune Watanabe, general manager of the laboratory, says that he sometimes has difficulty explaining to production managers the relevance of what is going on in Tsukuba. But the upper echelons of NEC are committed to backing basic research.

Until recently, NEC had only one research laboratory, in Kawasaki, not far from Tokyo. But now the company has five. A laboratory adjacent to a semiconductor plant in nearby Sagami-hara acts as a 'liaison office' between the research and development sections of NEC and production. The new laboratory in Tsukuba and another in Princeton 'New Jersey' in the United States, on the other hand, are devoted to much more basic research. A fifth laboratory for computer technology, the NEC Kansai C&C Research Laboratory, has just opened in Osaka.

Behind these moves to establish separate laboratories lies a deliberate strategy to create the freedom for more creative research by physically separating the people involved in production from those involved in basic research. Ironically, this is quite the opposite of what is going on in the United Kingdom and other Western countries, where government policy aims at bringing academics into closer contact with industry.

D.S.

ROMANIAN SCIENCE

Famous foreign faces

AMONG the wave of new companies coming to Tsukuba are several well-known foreign names, including the giant pharmaceutical and chemical concerns Glaxo, ICI and Upjohn, and the US semiconductor manufacturers Intel and Texas Instruments.

Upjohn's huge laboratory with 170 researchers was completed in 1988 at a cost of ¥14,000 million (\$87 million). A spokesman says that the decision to set up in Tsukuba is part of a global strategy to have research centres in the three key markets of the world: the United States, Japan and Europe.

Glaxo gives similar reasons for spending ¥20,000 million on a laboratory that will open in 1992 with a planned staff of 250 researchers.

At present, about 70 per cent of Upjohn's efforts are devoted to the clinical development of drugs for the Japanese market, which is second in size only to that of the United States.

But in the future the institute intends to devote more effort to the development of new drugs for markets throughout the world by "tapping into the rapidly expanding resource of Japanese pharmaceutical science".

Similarly, Texas Instruments intends to expend much of its efforts on basic research when it opens an institute for semiconductor research next year with an initial complement of 100 to 150 researchers drawn mainly from Japan. D.S.

industrial users, for example, and there are over 80 'research clubs' many of which are centred on Tsukuba University.

In addition, the local Ibaraki prefectural government, the Japan Development Bank, and 74 companies last year established the Tsukuba Research Support Centre, with a fund of ¥28,000 million (\$175 million) to provide companies and researchers with information on the latest developments in the science city. But while information networks may be well-developed, Tsukuba's transportation system remains grim. The only direct route to Tokyo is by bus which takes just over an hour, provided there are no traffic jams and a seat is available. Local government officials promise a direct train line to Tokyo within the next ten years, and there are also plans to build a highway in a huge loop around Tokyo passing near Tsukuba and joining the ends of the yet-to-be-built Tokyo Bay bridge.

If either plan is realized, Tsukuba can expect another surge in development and as one researcher put it "I can sell my house and retire on the proceeds".

David Swinbanks

Back to the future

Munich

IN a bid to encourage the rebirth of culture and science, researchers and professional people have reestablished Romania's Atheneum Society, a scientific academy banned in 1947 by the Communists.

The refounding of the society is important because the Romanian Academy is now discredited. The founders believe that only by starting a new and separate organization can science and culture be restored to their proper place in Romanian society. But vice president Ion Mircea Enescu, an architect, emphasizes that the society is not intended to replace the academy.

The original Atheneum Society was founded in 1865 as a cultural and learned society with 33 sections, from biology, law, fine arts and medicine to architecture and "fundamental science" — physics, mathematics, geometry and philosophy. It will be reconstituted in the same form, with the number of members fixed at 151.

Significantly, the Culture Ministry in the interim government of Ion Iliescu has offered its support to the society, but the amount is not yet clear. The society hopes to receive support through 'subscriptions' sold in Romania and abroad.

AIDS

One finger in the Romanian dike

Munich

A NEW report from the Romanian government suggests that efforts by it and foreign organizations are failing to stop the spread of AIDS in Romania. The report reveals that there were 478 cases of clinical AIDS in Romania as at 1 May, far more than the earlier estimates of 100 cases.

Even the new figure may be a drastic underestimate of the true situation. It is not known how many people are infected with HIV, the virus that causes AIDS.

Most alarming, 428 of the cases were found in children less than 13 years old. Of these 424 are under four. The now discontinued practice of giving 'microtransfusions' of blood to malnourished infants, many of them orphans, is thought to be behind the explosive spread of AIDS (see *Nature* 343, 579; 1990).

The blood supply, and the needles and syringes used, were contaminated by the AIDS virus. In a random nationwide sample of blood taken from children in orphanages, nearly ten per cent of the children tested positive for HIV, although the World Health Organization (WHO), based in Geneva, warns that this number might be high because of double reporting.

Testing of blood donors (there are an estimated 600,000 in Romania annually) and of blood already stored in blood banks

"We want only professionals" for the new society, says Enescu. The society, which has already nominated 100 members, will select only people considered politically "beyond reproach", he says. A very few were in the Communist Party but none were "professionally involved". None of the current academy members have been asked to join.

Another difference is the age of the members. Enescu says that the Atheneum Society will nominate only people who are still active in their fields. The current nominees are between the ages of 45 and 70. The average age of academy members is over 75 years.

At first, the society will sponsor public lectures and conferences on topics of current relevance such as ecology and the "philosophy of architecture and construction in Romania", says Enescu. The society may also advise the government in some capacity, though "we don't yet know the mechanism for this", he says.

Enescu is hopeful that the society will receive political as well as financial support from the Iliescu government, even though he pronounces the 20 May election that swept Iliescu to the presidency "a fake".

Steven Dickman

is not being carried out systematically, partly because of a shortage of testing equipment. Disorganization in the Health Ministry is also to blame.

Foreign relief organizations have delivered modern test equipment as well as thousands of sterile syringes and needles for injections and transfusions. The French organization *Medicins du Monde* alone has provided three ELISA (enzyme-linked immunosorbent assay) readers and thousands of test kits to laboratories in Bucharest for AIDS testing. But the Health Ministry has begun testing blood only in Bucharest and the three hard-hit eastern cities of Constanta, Iasi and Giurgiu. At least 20–25 laboratories including readers and kits will be required before blood screening can be comprehensive.

Romania's AIDS troubles may have only begun. Testing blood is a relatively simple job, says epidemiologist David Heymann of WHO, compared to the monumental task of educating the population about AIDS. According to Heymann, the Romanian press is just beginning to tackle the job. The regime of late dictator Nicolae Ceaucescu, which was overthrown in a bloody revolution in December 1989, had always claimed that AIDS did not exist in Romania and even forbade physicians to discuss it publicly. Steven Dickman

Problems of tenure in Japan

SIR—A recent Commentary (*Nature* 340, 337; 1989) noted that “in practice it is very difficult if not outright impossible for a foreign national to get tenure” in Japan, but discussed neither the cause of this difficulty nor how it could be ameliorated. The small number of tenured foreign faculty and the outright refusal by most universities to grant tenure to foreign faculty are particularly incongruous in the light of a recent proposal by Tokyo University “to create a science faculty which is also an internationally connected research community” (see *Nature* 338, 99; 1989). One of the goals of Tokyo University’s reform plan is to create an “intellectual climate that will be attractive to people from overseas”.

However, the “people from overseas” will be temporary employees, mostly at the postdoctoral level, and will not be encouraged to apply for permanent positions. It is questionable whether such a plan should be characterized as “internationalization”. Even on a purely pragmatic level, why does anyone think top foreign scientists will be interested in working in temporary posts in a far-away country with an unfamiliar language, where the only available career path is getting the boot?

I have been an associate professor in the department of geophysics of the faculty of science of Tokyo University since August 1984. My being hired as a tenured faculty member was due to the extraordinary efforts of Professor M. Kumazawa and his colleagues, who were committed to a policy of non-discrimination in hiring foreign faculty. My experience has been, on the whole, extremely positive from both a personal and scientific point of view.

Unfortunately, however, almost all foreign faculty members (that is, regular academic staff, excluding research associates) of Japanese national universities (including, hereafter, research institutes and joint-use institutes of *Monbusho*, the Ministry of Education, Science and Culture) are appointed for terms of one to ten years, with the possibility (but not a guarantee) of reappointment, whereas their Japanese colleagues are tenured. Temporary foreign faculty members, who are only nationally equal to their tenured Japanese colleagues, cannot readily become part of the Japanese scientific structure. Students, who require professorial assistance in finding employment, are unlikely to regard a temporary foreign professor as a desirable adviser. Finally, the temporary foreign staff cannot safely dispute the views of the tenured Japanese faculty who will decide on their contract renewals.

As of 1 July 1989 (statistics provided by

Monbusho) Japanese national universities employed a total of seven tenured foreign faculty members, but there were also 113 untenured foreign faculty members, with an average contract term of 2 years and 10 months. The following table summarizes the most recently available employment statistics. Note that all Japanese nationals are tenured, while almost all foreign faculty are temporary.

Japanese law was changed in 1982 to permit the employment of foreign nationals as faculty members at Japanese national

Rank	Foreign nationals		Japanese nationals	
	Tenured	Temporary	Tenured	Temporary
Professor	2	21	16,132	0
Assoc. profs.	4	52	14,276	0
Lecturers	1	40	5,216	0
Totals	7	113	35,624	0

universities. The regulations issued by *Monbusho* to implement the new law state in part (my translation) that “in principle foreign faculty should be hired on a temporary basis, but, in unavoidable circumstances”, the granting of tenure to foreign faculty “is not illegal”. (The cognizant governing body of each university makes the final decision on what constitutes an unavoidable circumstances.) However, only five universities (Tohoku, Tokyo, Niigata, Kyushu and the National Laboratory for High Energy Physics) use this clause to employ tenured foreign faculty members.

The extremely small number of tenured foreign faculty members at Japanese national universities is thus primarily due to the rules and personnel decisions of individual national universities, rather than to Japanese law or *Monbusho* regulations. Why do almost all Japanese universities refuse to grant tenure to foreign faculty? Why do the handful of universities that will grant tenure to foreign faculty do so only in a small number of cases? And what, if anything, can be done to improve the situation?

First, no matter how much the administrators of Japanese national universities talk about internationalization, they have not truly grasped the notion that the development of an open university requires fully equal treatment of foreign and Japanese faculty. Second, because there is excessive concern over what would happen if a tenured foreign faculty member proves unsatisfactory, administrators prefer to retain control by hiring foreign faculty on a temporary basis. This eliminates the risks (both real and perceived) of hiring tenured foreign faculty, but also eliminates the possible benefits, which, it seems to me, far outweigh the risks. In any case, problems involving tenured Japanese nationals are by no means unknown.

Finally, it should be noted that Japanese universities are not fully open even on a national level. For example, in 1987, 86.8 per cent of the academic staff of Tokyo University were graduates of Tokyo University (*AERA*, No. 15, p. 15; 1989); the existence of such inbreeding is inconceivable at, say, any of the US universities. The failure of Japanese national universities to grant tenure to foreign faculty is only one aspect, then, of their general lack of openness, a theme I have discussed at length elsewhere (*Kagaku* 56, 566; 1986). Further opening of Japanese universities to foreign faculty should also lead to a great increase in the movement of personnel between universities within Japan.

Unless the present discriminatory tenure policies are scrapped, I believe that the internationalization of Japanese higher education is doomed to fail, to the detriment of all concerned.

ROBERT J. GELLER

Faculty of Science,
Tokyo University,
Tokyo, Japan

Views of the staff

SIR—In “The museum director’s view . . .” (*Nature* 345, 198; 1990) the director of the Natural History Museum, Dr Neil Chalmers, states “we have chosen to concentrate on those areas of taxonomic research where we already have the greatest strength. . . . We have identified six such areas . . .”. His use of “we” in this context could imply that the museum scientists took some corporate decision to alter the direction of research at the Natural History Museum.

Nothing could be further from the truth. The decisions were taken by a group of seven; the director, associate director and five heads of science departments, few of whom are involved in collection based research. The remainder of the museum’s 300 scientists were not consulted.

Such is the sense of outrage at the museum that our senior scientists have volunteered to set up a committee, working with their trade union, IPMS, to defend science at the Natural History Museum, the museum staff are experts in the field of taxonomy, which the director is not. They are very clear about where the science priorities lie, and it is tragic that the director excluded this body of expertise from development of the corporate plan. We now face a protracted campaign to reverse the damaging decisions made in the corporate plan — a wholly avoidable situation.

PENNY WHEATCROFT
(Branch Chair)

Institution of Professional
Managers and Specialists,
British Museum (Natural History) Branch,
Cromwell Road, London SW7 5BD, UK

Understanding gel electrophoresis

Pulsed-field electrophoresis, the newest version of a standard technique, has become widely used for the sieving of large DNA molecules long in advance of an understanding of exactly how it yields useful results.

ELECTROPHORESIS may be one of the most widely used techniques of molecular biology, but its practice is still a craft. Some researchers, or their technicians, have acquired reputations for particular skill in the preparation of gels, or in the interpretation of the records won this way. And while the principles of gel electrophoresis are simple enough — the rate at which protein or DNA molecules are dragged through the aqueous interstitial network running through an inert polymer skeleton by an electric field depends on the length of the molecules, often a good proxy for mass — the limitations of the technique have usually been discovered the hard way (by experiment) before they have been accounted for by calculation.

That is why major innovations in gel electrophoresis have usually also arisen empirically. That is notably the case with the latest development, that of pulsed-field electrophoresis, in which the direction of the forcing electric field alternates at regular intervals between a forward direction and one that may be diametrically opposite or even include a substantial orthogonal component. (The technique, due to David C. Schwartz and Charles R. Cantor, is described in *Cell* **37**, 67; 1984.) The technique has made it possible to separate larger molecules than are accessible when the electric field has a single direction — up to roughly 2 million nucleotide bases, compared with the limit of about 30,000 bases for the conventional technique. But why?

Hand-waving explanations are plentiful, but throw little light on how to optimize the new technique for the separation of molecules of particular size and conformation. Because the applied electric field (usually no greater than a few volts per centimetre) is invariably in the plane of the thin gel, the problem is that of the migration of long molecules in two dimensions. Most of the medium is aqueous, with the polymer skeleton (polyacrylamide or agarose) represented as a random scattering of physical obstacles which long molecules cannot penetrate but on which they can become entangled. Because the medium is aqueous, DNA molecules therein are uniformly charged along their length. In an electric field, they may be expected to drift in that direction. But at what rate?

This is where the fun begins. Suppose that, against all the odds, a DNA molecule is already stretched in the direction of the

electric field, which in the simplest case is constant. The molecule has a leading end (the 'head') and a trailing end (the 'tail'). The head must find a path through the solid obstacles randomly presented by the gel, and the remainder of the molecule must, snakelike, follow (whence the description of the movement as 'reptation').

The simplest view of what happens is to suppose that the molecule moves within a tube, more or less in the direction of the field — that defined by the path sought out by the head of the molecule. But there is an obvious difficulty. The unthinking head, subject to unavoidable thermal fluctuations, will hunt in one direction and another for a path through the obstacles in its way, but the path for the remainder of the molecule is already defined, so that the speed at which the tail end sections move behind the head will be determined only by the time taken for thermal fluctuations, as biased by the electric field, to carry them along the tube.

The mobility of the molecule as a whole will be determined by the likelihood that its successive parts will move in the direction of the electrical bias. This is essentially a random walking problem, so that the mobility is proportional to the square root of the length (or mass) of the molecule. So the resolution of electrophoresis should be inversely proportional to the length of the molecules whose separation is attempted.

Sadly, there are complications. For one, even on the assumption that the migration of the molecule represents motion along some notional tube, its trailing sections will travel faster than the head, so that a molecule that began as a stretched-out string will become a more tightly knit conglomerate. The longer the molecule or the stronger the electric field, the more quickly this will happen, when the mobility will be determined by the resistance to the passage of the tangle, whose dimensions will vary only slowly (say, as the one-third power) with length. The mobility of large molecules will be anomalously greater than expected, and sufficiently large molecules will be indistinguishable from each other by their mobility.

Simple (indeed, over-simple) as it is, this picture gives a hint of how pulsed-field electrophoresis may help with the resolution of larger molecules. If the field is reversed, the trailing tangle will be partly

unravelling, restoring molecular mobility at least for a time. Obviously, the reversed field must not last for longer than the direct field, or there would be no resolution of even short molecules unaffected by the tangling of trailing molecular sections.

Unfortunately, this picture is a drastic simplification, which takes no account of how molecules become entangled with the gel lattice. So, for the past five years, there has been a steady sprinkling of attempts to model the movement of polymer chains through gel lattices, supposed to consist of regular arrays of obstacles. The general idea is to represent the molecule by a notional string of beads, each carrying an electric charge, which may be connected inelastically or (more cleverly) by elastic springs. The problem is to estimate the rate at which such a structure moves through the lattice of obstructions.

J. M. Deutsch and T. L. Madden were among the first to show that there may be a curious and unexpected oscillation of a normally migrating polymer chain between the extended and tangled structures (*J. Chem. Phys.* **90**, 2476; 1989). Now, Monica Olvera de la Cruz, Dilip Gersappe and Edward O. Shaffer from Northwestern University, Illinois, have described an even more elaborate calculation (using a Cray machine) that throws a little more light on what happens in pulsed-field electrophoresis (*Phys. Rev. Lett.* **64**, 2324; 1990).

The oscillation between the bunched and extended configurations shows up clearly in calculations of the instantaneous radius of gyration (root mean square radius) of the migrating molecules. The beneficial effects of intervening pulses of electric field in a transverse direction show up both in calculations of molecular mobility and in snapshots of the simulations: a transverse field opens up tangled structures, assisting their migration in the principal direction (which nicely confirms the pictures reported last year by Schwarz and Koval in *Nature* **338**, 520; 1989). It is even possible qualitatively to relate the relative duration of the direct and transverse fields to the length of the large molecules for which the electrophoretic resolution will be greatest. Yet there is a long way to go before the problem can be said to be neatly tidied up — by which time, no doubt, there will be other innovations to account for.

John Maddox

Vesicular consumption

Graham Warren

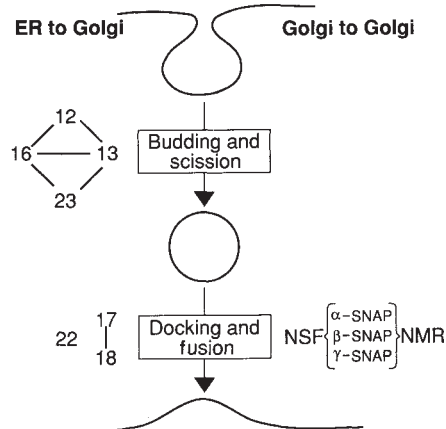
TRAFFIC of proteins between intracellular membranes is discontinuous, the process being mediated by coated vesicles which bud from one membrane and fuse with the next membrane in the pathway¹. Budding from the plasma membrane is mediated by adaptors which link clathrin coat subunits to the cytoplasmic tails of receptors to be internalized². Rearrangement of the clathrin lattice causes invagination to form the vesicle, which is released by scission and then uncoated in preparation for fusion with endosomes³⁻⁵. Fusion is best understood at the level of Golgi transport⁶, and although the mechanism is still obscure, two of the proteins involved have been purified. The first, NSF (for *N*-ethyl maleimide-sensitive factor)⁷, is involved in fusion after the uncoating of the bound transport vesicle⁸. The second, factor B (ref. 9) (renamed POP (ref. 10), for protein operating in pre-fusion), is involved at a step just before bilayer fusion. Subsequently, NSF was shown to be involved in other vesicle-mediated steps^{11,12}, implying that it is part of a general fusion machine. Three recent studies — by Kaiser and Schekman¹³, Clary and Rothman¹⁴ and Clary *et al.*¹⁵ — together identify another component of this machine using genetic and biochemical approaches.

The genetic experiments involved re-examination of yeast secretory (*sec*) temperature-sensitive mutants that prevent transport from the endoplasmic reticulum (ER) to the Golgi apparatus at the restrictive temperature. By definition, all accumulate ER, but earlier work had shown that some also accumulate small vesicles that seemed to be derived from the ER^{16,17}. Using permanganate fixation to highlight these vesicles so that they could be counted, Kaiser and Schekman¹³ have now shown that the mutants can be divided into two classes (see figure). The first (*sec12*, *sec13*, *sec16* and *sec23*) cannot produce vesicles at the restrictive temperature; the second (*sec17*, *sec18* and *sec22*) cannot consume them, so thousands accumulate in each cell. This classification predicts that double mutants, constructed from a member from each class, would be unable to produce vesicles — this was observed in all cases.

Lethal phenotype

Other double mutants, constructed from members of the same class, mostly yielded a lethal phenotype, a phenomenon known as synthetic lethality. One possible explanation for this is that although the mutant protein changes to a functional conformation at the permissive temperature, the overall structure still differs from the nat-

ive protein. This will not matter if the protein with which it normally interacts is native, but if it, too, is mutant, and also differs structurally at the permissive temperature, then the combined differences may clash to give a completely nonfunctional, lethal complex. The absence of lethality does not eliminate the possibility of protein-protein interaction but its presence strongly suggests it. For the *sec* mutants involved in the production of vesicles, five out of the six possible pairwise combinations are lethal. This sug-



The two classes of *SEC* gene involved in ER-to-Golgi transport are shown on the left, the connecting lines showing the lethal pairwise combinations¹³. The proteins involved in fusion of Golgi transport vesicles are shown on the right. *SEC18* is NSF (ref. 18), *SEC17* is α -SNAP (ref. 15). NMR, NSF membrane receptor¹⁹ which has still to be properly characterized.

gests considerable interaction, which might be expected if the encoded proteins bear any resemblance to the components involved in the budding of endocytic vesicles. By contrast, of the three *sec* mutants involved in vesicle consumption, only one pair is lethal, *sec17* and *sec18*. Because *SEC18* encodes NSF (ref. 18), *SEC17* should encode a protein that interacts with NSF. This protein has now been identified and its interaction with NSF confirmed in an elegant series of experiments^{14,15}.

NSF is a peripheral membrane protein which is isolated in a soluble form⁷ and will bind to Golgi membranes only in the presence of a cytosolic fraction termed SNAP (soluble NSF attachment proteins)¹⁰. By using transport activity as the assay, Clary and Rothman¹⁴ purified to homogeneity three proteins that operate at the same stage of transport as NSF. Clary *et al.*¹⁵ went on to show that these proteins have SNAP activity. The purified SNAPs (α , β and γ) each bind NSF to Golgi membranes and work independently of one another. The amount of SNAP

needed to bind NSF to membranes is very similar to the amount needed for vesicle transport, providing the best evidence so far that NSF must be bound to Golgi membranes in order to function. Further experiments showed that SNAP proteins alone would bind NSF and therefore that they link the NSF directly to the Golgi membrane receptor (NMR; see figure).

Yeast cytosol was demonstrated to have SNAP activity enabling *sec* mutants to be screened. The *sec17* mutant was inactive and activity could be restored by the addition of α -SNAP but not β - or γ -SNAP. This result both complements and confirms those obtained by Kaiser and Schekman. It also predicts that the yeast genes encoding β - and γ -SNAP should yield *sec* mutants that accumulate vesicles.

Binding proteins

What then is the role of each of the three SNAPs? They could be viewed as members of a large family of binding proteins, each operating at a different step in the pathways of membrane traffic, binding NSF to target membranes. As the authors of the new papers argue, however, this does not explain the complementation of *sec17* by α -SNAP. The *sec17* mutant interrupts transport from the ER to Golgi, yet α -SNAP was identified by its ability to transport vesicles from *cis* to medial cisternae.

Therefore, α -SNAP operates at more than one step. Instead, as the authors point out, the proteins should be viewed as comprising a particle containing NSF, α -, β - and γ -SNAPs. It is then necessary to explain why any one of the SNAPs alone can substitute for the other two in transport and binding assays. In the binding assays, one can imagine that any one of the SNAPs might suffice to bind NSF to Golgi membranes. In the transport assay, however, all should be needed if the particle is to function. The authors argue that washed Golgi membranes may still contain residual SNAPs in amounts sufficient

- Palade, G. *Science* **189**, 347–358 (1975).
- Pearse, B.M.F. *EMBO J.* **7**, 3331–3336 (1988).
- Heuser, J. & Evans, L. *J. Cell Biol.* **84**, 560–583 (1980).
- Smythe, E. *et al. J. Cell Biol.* **108**, 843–853 (1989).
- Rothman, J.E. & Schmid, S.L. *Cell* **46**, 5–9 (1985).
- Rothman, J.E. & Orci, L. *FASEB J.* **4**, 1460–1468 (1990).
- Block, M.R. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 7852–7856 (1988).
- Malhotra, V. *et al. Cell* **54**, 221–227 (1988).
- Wattenberg, B.W. & Rothman, J.E. *J. Biol. Chem.* **261**, 2208–2213 (1986).
- Wattenberg, B.W. *et al. J. Cell Biol.* **110**, 947–954 (1990).
- Beckers, C. J. M. *et al. Nature* **339**, 397–398 (1989).
- Diaz, R. *et al. Nature* **339**, 398–400 (1989).
- Kaiser, C.A. & Schekman, R. *Cell* **61**, 723–733 (1990).
- Clary, D.O. & Rothman, J.E. *J. Biol. Chem.* (in the press).
- Clary, D.O., Griff, I.C. & Rothman, J.E. *Cell* **61**, 709–721 (1990).
- Novick, P. *et al. Cell* **25**, 461–469 (1981).
- Esmon, B. *et al. Cell* **25**, 451–460 (1981).
- Wilson, D.W. *et al. Nature* **339**, 355–360 (1989).
- Weidman, P. J. *et al. J. Cell Biol.* **108**, 1589–1596 (1989).
- Goud, B. *et al. Cell* **53**, 753–768 (1988).
- Baker, D. *et al. Cell* **54**, 335–344 (1988).
- Ruohola, H. *et al. J. Cell Biol.* **107**, 1465–1476 (1988).

to complement the exogenously added SNAP. But if each purified SNAP can be complemented, all must be present on the membrane, so why would any exogenous SNAPS be needed? Further work is required both to resolve this question and to define more precisely the involvement of each of the components in this complex fusion machine.

The study of membrane traffic has in the past been dominated by morphological and biochemical techniques^{1,6}. The *sec* mutants were first described ten years ago¹⁶, but, with one notable exception²⁰, they have not in themselves given insight into the mechanism of action of the molecules they encode. One reason has been the lack of cell-free assays that mimic steps on the secretory pathway, although considerable progress has been made in the past few years^{21,22}. Another and more serious reason is the absence of good morphological techniques for studying the

pathways of membrane traffic in yeast. Here is an area where microscopists willing to make the effort will reap the reward of raising the morphological quality to the level long enjoyed by animal cell microscopists. Kaiser and Schekman show what can be done when one combines yeast genetics and morphology, but the fact that they had to use such a drastic staining technique, which destroyed much of the cell's morphology and reduced the vesicles they were measuring to amorphous, metallic blobs, serves to emphasize the effort that still needs to be applied in this area. In the absence of such effort, morphological approaches to this problem will continue to depend upon hybrid systems in which animal membranes are complemented by yeast components. □

Graham Warren is at the Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK.

EARTHQUAKES

Parkfield slowing down

William D. Stuart

EARTHQUAKE precursors are so elusive that their detection requires dense arrays of sensitive instruments close to the area of earthquake rupture. But to select an array site one first has to predict coarsely by other means where the earthquake will be. For the San Andreas fault near Parkfield, California, a coarse prediction has been provided by the roughly periodic occurrence of similar moderate earthquakes (magnitude 5.5–6) in 1966, 1934, 1922, 1901, 1881 and 1857 (ref. 1). From these dates alone the expected time of the next event was 1988 with an uncertainty of ± 7 years. Using data from a dense seismometer array operated by the US Geological Survey in anticipation of the next earthquake, Wyss *et al.*² on page 426 of this issue report a rate decrease of small-magnitude seismicity starting in 1986. They interpret this quiescence as a precursor and estimate that the next moderate earthquake will occur in March 1991 ± 1 year. Thus, a long-term prediction has enabled an intermediate-term one to be made.

Wyss *et al.* estimate the time of the impending earthquake by comparison with six previous periods of quiescence that preceded a main-shock. The case histories used were for earthquakes of magnitudes 3.6–8 in California, Hawaii and Alaska. Curiously, the duration of the period of quiescence for these cases varies from 1.5 to 3.8 years, suggesting a weak dependence on

the earthquake magnitude. But the durations and earthquake magnitude are positively correlated.

Having so few case histories and lacking a physical theory, Wyss *et al.* acknowledge that the uncertainty of ± 1 yr is poorly constrained. If one accepts the case histories, the four-year lapse since the onset of quiescence is actually more consistent with a magnitude 7 earthquake than a magnitude 5.5–6 earthquake. The San Andreas fault just southeast of Parkfield is

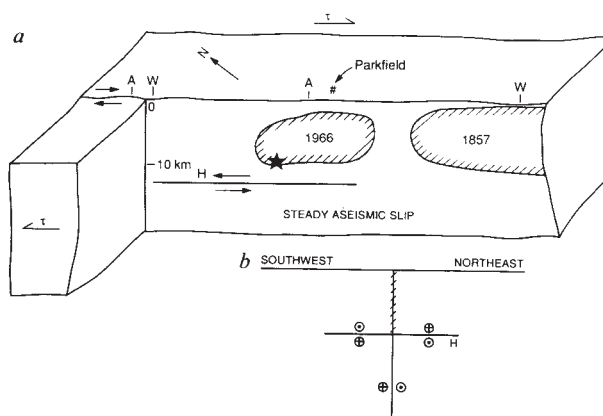
a likely place for a magnitude 7 event³. That fault section has been locked since the magnitude 8.3 Fort Tejon earthquake in 1857, and the amount of slip now available for release during an earthquake is 4–5 m.

The quiescence in the Parkfield earthquake catalogue was also noticed by Aviles and Valdes⁴. They found that the occurrence rate of earthquakes of magnitude greater than 2 decreased over a crustal volume starting about half way into the rupture area of the 1966 earthquake and extending 30 km northwest along the fault (see part *a* of figure). The crustal volume studied by Wyss *et al.*, on the other hand, is 54 km long and essentially encloses both the volume of Aviles and Valdes and the entire 1966 rupture area (since healed). Both groups agree that the strongest quiescence is in the smaller volume, but it is unclear why the strongest quiescence should enclose just the north end of the 1966 rupture and not all of it.

Reasenber and Matthews^{5,6} formulated a precise statistical procedure for a quiescence search in an earthquake catalogue and dispute some of the quiescences claimed by Wyss *et al.*². In particular, they argue that the claimed statistical significance of quiescences is too high, and that when the correct statistic is used some rate changes of seismicity that seem to be visually significant are in fact not. It may be that only certain fault geometries or slip modes have quiescences.

The case for the Parkfield quiescence is considerably strengthened by geodetic distance measurements made by a two-colour laser-ranging instrument. In a companion paper on page 428 of this issue, Wyss *et al.*⁷ notice that the rate of shortening of two geodetic survey lines crossing the San Andreas fault decreased by about 20 per cent within a few months of the start of the period of quiescence. So there is good evidence of either some physical change in the fault or an alteration of loading rate. One explanation suggested by Wyss *et al.*⁷ is an increase in the fraction of the fault zone that is locked. Overall, the fault would become stronger and consequently deform at a lower rate. A lower deformation rate implies decreased production of background seismicity and slower motion of bench marks at the ends of geodetic lines.

Another model that comes to mind is a simple revision of the generally accepted model for stressing of the fault near Parkfield. In the current model (see figure) the top 10–20 km of the San Andreas fault is locked from the 1966 rupture south. Elsewhere the fault



a, Cutaway view of San Andreas fault at Parkfield. Regional stress tending to deform the Parkfield area is shown by τ . Paired arrows show sense of fault slip. Hatching outlines currently locked fault patches. '1966' marks the patch expected to fail again soon in a moderate earthquake (star is the focus of the 1966 earthquake); '1857' marks the tip of the patch that failed to make the $M = 8.3$ Fort Tejon earthquake. A and W mark the edges of the crustal volumes used by Aviles and Valdes⁴ and Wyss *et al.*² to detect seismic quiescence. H marks the position of the conjectured horizontal fault responsible for both quiescence and the rate decrease of geodetic lines. *b*, End view of fault showing relative displacements for fault slip. The dot and plus sign indicate motion out of and into paper, respectively.

slides aseismically, loading the locked patches towards failure. Moderate earthquakes, such as the 1966 event, occur during failure of the left patch shown. Otherwise, the stress increase on and near locked patches generates background seismicity and deformation of the ground surface.

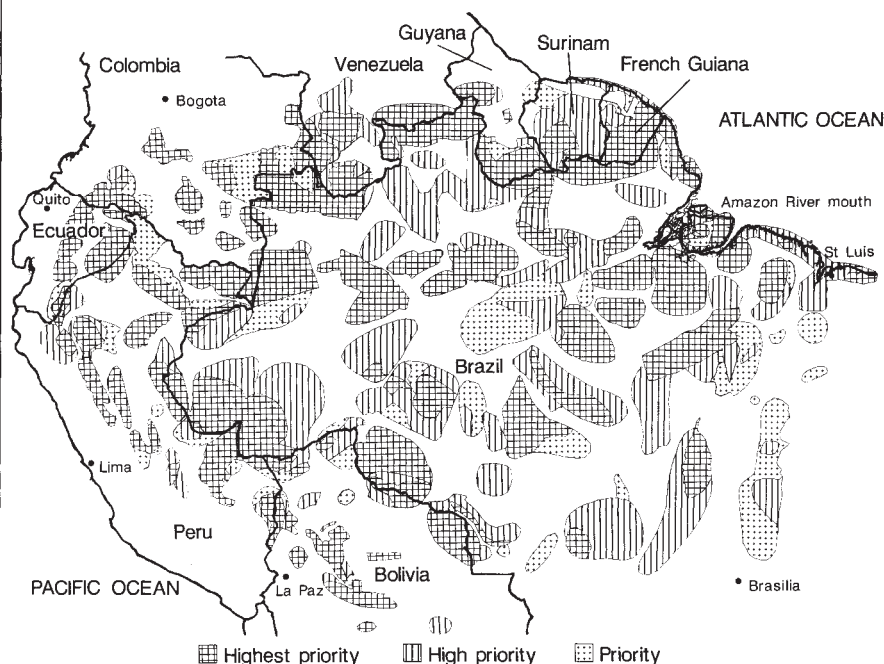
In the revised model, seismicity and geodetic-rate decreases are caused by slip on a sub-horizontal fault (H in figure) beneath the locked patches. The same deep slip on the San Andreas fault that loads the locked patches also loads a horizontal plane at H, and if the plane's yield stress is a little less than that of the over-head patch, then the horizontal fault will slip first. Slip on the horizontal fault will be late in the seismic cycle and will decrease the loading rate on the vertical fault and adjacent crust. The postulated fault may be part of the sub-horizontal detachment fault associated with the magnitude 6.7 Coalinga earthquake in 1983, 40 km northeast of Parkfield. In short, seismicity and crustal motion slow down late in the earthquake cycle because the detachment fault, acting as a buffer, comes to life.

The reason for seeking precursory anomalies is, of course, to make short- or intermediate-term earthquake predictions. The advantage of such predictions is that they may be more precise than long-term ones. Long-term predictions (more than a few years), such as those derived from earthquake-recurrence intervals, usually have uncertainties of many years. This is true even when long-term predictions are refined using conditional probabilities to allow for the time elapsed with no earthquakes³. As Savage⁸ puts it, the probabilities are uninformative because the acceptable range of probability values is so great.

The empirical approach of Wyss *et al.*^{2,7} is one way to make an intermediate-term prediction from observations. But their work also places within reach a mechanical model connecting background seismicity, crustal deformation and a forthcoming mainshock. A detailed and accurate model might therefore be adapted to forecast earthquakes just as atmospheric models are used to forecast the weather. □

William D. Stuart is at the US Geological Survey, Pasadena, California 91106, USA.

Consensus for conservation



THE map reproduced here, the fruit of a recent interdisciplinary meeting*, identifies the areas of biological priority for conservation in Amazonia, graded into three levels. The total area shaded covers about 60 per cent of the Amazon basin and the areas with square hatching are of greatest biological importance. The map represents the consensus of some 100 biologists, physical scientists and conservation planners — it will be an essential tool in planning the location of biological reserves, national parks, extractivist reserves and other conservation areas in Amazonia.

Scientists at the meeting began by working in small specialist groups: plant systematics, plant ecology, mammals, ornithology, herpetology, ichthyology, entomology, geomorphology and climate, and units of conservation. The seven biological groups each produced thematic maps to justify their selection of priority areas. This was followed by separate meetings of botanists and zoologists to pool their data and consolidate their chosen areas on maps for botany and for zoology; as it turned out the areas selected by the different groups broadly corresponded. At the same time, the geomorphologists and climatologists produced a map identifying the most fragile soils and the ecosystems requiring the greatest protection. The conservation unit's group discussed policy and other factors bearing on the preservation, conservation and management of priority areas.

The last step was the fusion of the 104

areas of biological importance for botany with the areas selected by zoologists. The final map shows 94 priority areas, evaluated on a five-point scale (eventually reduced to three). Areas with the maximum overlap between disciplines were given the highest priority.

Much on-site work remains to be done by conservation organizations of the nine countries concerned, to select the precise areas suitable for different categories of reserves, managed forests and regions for the indigenous population. The shaded parts of the map include some deforested and inhabited areas and site visits are necessary to locate suitable reserves. But the biologists have provided their data, and now it is the turn of conservation planners to choose appropriate reserves, politicians to make them a reality and international assistance organizations to help finance them.

What of the white areas? Are these the regions in which deforestation should be allowed to proceed? They are indeed probably of less importance in terms of species diversity, but the state of inventory of the Amazon region is still so poor that it would be unwise to advocate development projects in these spaces that, in many cases, represent some of the least explored parts of Amazonia. These are areas that need to be chosen for further inventory, rather than for deforestation.

Ghilleen T. Prance

Ghilleen T. Prance is at the Royal Botanic Gardens, Kew, Richmond, Surrey TW9 3AB, UK.

*Workshop 90, Manaus, Brazil, 10–20 January 1990.

1. Bakun, W.H. & Lindh, A.G. *Science* **229**, 619–624 (1985).
2. Wyss, M., Bodin, P. & Habermann, R.E. *Nature* **345**, 426–428 (1990).
3. *US Geol. Surv. Open-File Rep.* 88–398, (USGS, 1988).
4. Aviles, C.A. & Valdes, C.M. *Eos* **70**, 1228 (1989).
5. Reasenber, P.A. & Matthews, M.V. *Pure appl. Geophys.* **126**, 373–406 (1988).
6. Matthews, M.V. & Reasenber, P.A. *Pure appl. Geophys.* **126**, 357–372 (1988).
7. Wyss, M., Slater, L. & Burford, R.O. *Nature* **345**, 428–431 (1990).
8. Savage, J.C. *Seismol. Res. Lett.* **61**, 43 (1990).

Agricultural renaissance in the high Andes

Warwick Bray

ONE of the saddest sights in Peru and Bolivia is the vast extent of farmland that was intensively cultivated in prehispanic times and is now all but abandoned. Between 50 per cent and 75 per cent of Inca terraces are no longer in use, and in some of the coastal valleys there was up to 40 per cent more irrigated land in pre-Conquest times than there is today (Clark Erickson, manuscript in preparation). Moreover, entire technologies have been forgotten: the *qochas* (circular field basins) of the high Andes, the sunken gardens in areas of high ground-water along the coast, and the systems of raised fields in waterlogged areas at all altitudes. The remains of these indigenous technologies are now being rediscovered by archaeologists and geographers, and the implications have forced government agronomists and foreign-aid agencies to take notice.

The most sophisticated of the current projects are located in the basin of Lake Titicaca, on the Peru–Bolivia border at about 3,800 m above sea level. Aerial photographs revealed more than 83,000 hectares of eroded and abandoned fields on the lakeshore plain in terrain nowadays used only as low-grade pasture. The zone adjacent to the lake is perennially marshy and is subject to flooding in the wet season. These problems are compounded by efflorescence of salts as a result of evapotranspiration, and by the risk of sudden frosts that can wipe out an entire harvest.

Investigations began in 1981–86 when Erickson excavated a group of raised fields near Huatta, on the Peruvian side of the lake^{1,2}. He showed that between 1000 BC and AD 400 the inhabitants of the region dug a series of parallel canals, and piled the earth between them to form long, low mounds about 1 m high, 4–10 m wide and 10–100 m long. Pollen from the field soils indicates that potatoes and quinoa (a high-altitude grain rich in protein) were the main crops. The canals provided moisture during periods of drought, and were also a source of green mulch and organic muds to fertilize the fields, thus eliminating the need for fallowing. The raised planting surface helped to prevent waterlogging of roots, and the canals served as traps for sediment (which could be scooped out and recycled) and also as heat sinks, warming up during the days and slowly releasing heat at night. Measurements show that the night temperature on raised fields in conditions of light frost was about 2 °C higher than that of control plots, and that the frost was

several hours shorter in duration. The effect was not merely to reduce frost damage, but also to extend the growing season.

In the Huatta region, construction of raised fields ceased by about AD 400, a date that coincides with the growth of the city of Tiwanaku in the southern (Bolivian) part of the basin. In the rural hinterland around Tiwanaku, Kolata and

destroyed harvests on conventional plots, losses of potatoes, barley, quinoa and vegetables on the raised fields were minimal.

Above all, as Erickson and his colleagues emphasize^{1,2}, raised-field technology requires no imported seed, chemical fertilizers or expensive machinery, is based on traditional crops and tools, and does not disrupt existing work patterns and social norms. In contrast, other competing schemes have been capital-intensive or designed to produce cash crops from which most small farmers have not benefited. By starting at the bottom, with the villagers themselves (rather than imposing from the top), by setting a

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

The shores of Lake Titicaca, in prehispanic times a site of intensive agriculture.

Ortloff are studying the agricultural landscape that provided food for one of the main prehistoric civilizations of the Andes. They show that the landscape of the Tiwanaku period (about AD 400–1100) was almost entirely man-made, with terraces on the hill slopes and systems of raised fields, interspersed with farms and townships, on the lakeshore plain^{3,4}. Using mathematical models to study water flow and the effects of fluctuating water levels on drainage, salinization and heat-storage capacity, they demonstrate that the landscape functioned as an integrated system. Wet-season rainfall in the hills was collected in modified natural gullies and transferred by means of aqueducts across the plain and into the lake. Canalizing runoff prevented deposition of unwanted sediment and, more importantly, controlled the water table by reducing infiltration of surface water into the groundwater reservoir.

But could this kind of system work today? The answer is 'yes'. On experimental raised fields, sustainable potato yields are twice those of the control plots, and often much more. In years when frost

practical example, and by initiating a training programme with simple manuals in Spanish and Quechua, the Huatta team has persuaded several local communities to build raised fields for themselves. Similar experiments have begun on a small scale in the very different conditions of coastal Ecuador⁵, and Colombia (with more than 500,000 hectares of abandoned raised fields in the Caribbean lowlands⁶) is also showing interest. □

Warwick Bray is at the Institute of Archaeology, University College London, 31–34 Gordon Square, London WC1H 0PY, UK, and at the Institute of Latin American Studies, University of London.

1. Erickson, C. L. & Candler, K. L. in *Fragile Lands of Latin America: Strategies for Sustainable Development* (ed. Browder, J. O.) 230–249 (Westview Press, Boulder, 1989).
2. Erickson, C. L. *Expedition* 30(3), 8–16 (1990).
3. Kolata, A. L. & Ortloff, C. J. *archaeol. Sci.* 16, 223–263 (1989).
4. Ortloff, C. R. & Kolata, A. L. *J. archaeol. Sci.* 16, 513–535 (1989).
5. Muse, M. & Quintero, F. in *Pre-Hispanic Agricultural Fields in the Andean Region* (eds Denevan, W. M., Mathewson, K. & Knapp, G.) 249–266 (British Archaeological Reports, International Series 359, Oxford, 1987).
6. Plazas, C., Falchetti, A. M., van der Hammen, T. & Botero, P. *Boletín del Museo del Oro* 55–88 (1988).

Electrons give a broader view

Archie Howie

THE dazzling success of optical holography cannot blind us to the fact that Gabor's proposal¹ in 1948 for atomic resolution by electron holography has so far been realized only partially, despite substantial advances in electron microscopy^{2,3}. Some new proposals⁴⁻⁶ for lensless holographic imaging of adsorbed surface structures using low-energy electrons may, however, have brought the prospect distinctly closer.

Electron diffraction patterns, Fresnel fringes and other wave-interference effects are easily demonstrated. But to record a hologram based on electron wavelengths, considerable magnification is needed, either through geometrical projection or with electron lenses still restricted by spherical aberration to an angular aperture of about 1° . Stray a.c. magnetic fields are often critical because they destroy interference effects. Moreover, the strong electron-atom interaction demands that a thin sample be used to minimize multiple elastic scattering and avoid the equivalent of the optical hologram of a chess set that is so highly reflective that each piece is illuminated mainly by light coming from the others by a variety of interference paths rather than directly from the source. Inelastic scattering is also likely, destroying any interference effects except in situations where the reference and scattered beams are

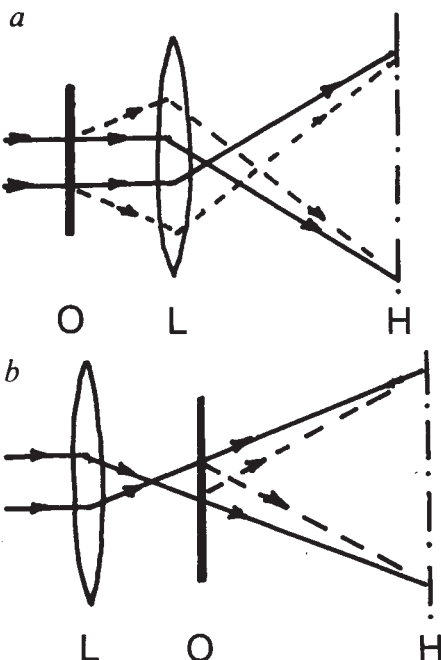


FIG. 1 Schematic illustrations of the use of a lens L to produce a transmission hologram H of a transparent object O by interference of a reference wave (continuous line) and a scattered wave (broken line). The standard arrangement for high-resolution electron microscopy is shown in a; Gabor's projection scheme¹ is shown in b.

involved in the same inelastic scattering event.

High-resolution transmission electron microscopy overcomes these problems by using as a reference wave that part of the incident beam that is transmitted without deflection through the specimen. The objective lens deflects the waves scattered by the specimen back to the axis where they interfere with the reference wave to form an on-axis, image plane hologram as shown in Fig. 1a. This hologram can be recorded on a photographic plate after further magnification. Thin samples behave like 'phase objects' (perfectly transparent but inducing a phase change) and a combination of defocus and spherical aberration is used to obtain amplitude contrast in the hologram, typically restricted to a range of object spacings between 2 and 0.15 nm. Critical extraneous a.c. magnetic fields are in the milligauss region. These image holograms can be viewed directly as electron micrographs without any reconstruction procedure and, because of the restricted lens aperture, look like two-dimensional projections of the three-dimensional object structure. The temptation to interpret them literally — and sometimes erroneously, because of complications from instrumental aberrations, multiple scattering or projection effects — has generally proved irresistible.

An electron-optical biprism incorporated in the microscope column allows more ambitious off-axis holographic imaging of phase-object-contrast detail from magnetic structures², p-n junctions⁷ and surface steps⁸ and may also enable the high-resolution limit to be pushed down to scales of 0.1 nm and below⁹. Gabor's geometrical magnification scheme¹ has been re-examined⁹ using a high-resolution scanning transmission microscope to form a focused source close to the object, as shown in Fig. 1b. In principle, the method will work if the instrumental aberrations are sufficiently well known to specify accurately the complex amplitude distribution in the source, but additional difficulties arise because of mechanical vibration and specimen drift. The restricted angular aperture of electron lenses, whether before or after the object, is also a continuing obstacle to genuine three-dimensional imaging at high resolution.

The proposal⁴ for surface holography elegantly achieves a coherent point source fixed to the sample by causing one of the surface atoms of the sample to emit either

low-energy photoelectrons or (following either photon or electron excitation) Auger electrons. As shown in Fig. 2, a hologram is then formed by simple geometrical projection of these electrons onto a hemispherical screen or energy-selective angular detector, where a reference wave coming directly from the source atom can interfere with waves that arrive after scattering on surrounding atoms. The source atoms can be surface adsorbate atoms, which need not be arranged periodically. Their individual contributions to the hologram intensity will superimpose pro-

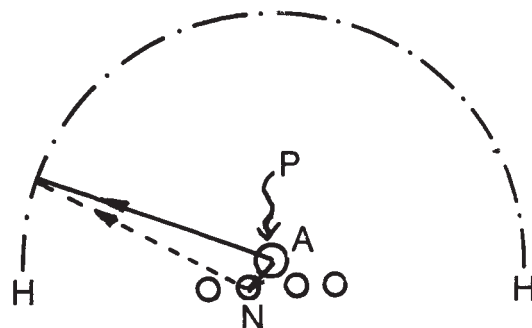


FIG. 2 Schematic illustration of the use of a source atom A (stimulated by a photon or high-energy electron P) to emit a spherical reference wave (continuous lines) interfering with weaker waves (broken line) scattered by neighbouring atoms N. A hologram H is thus formed on a remote spherical screen.

vided that they each sit on a site with the surrounding atoms in the same angular positions. Reconstruction would involve illuminating the hologram with a coherent converging spherical wave, but this optical problem can be solved numerically by Kirchoff's method. In a numerical simulation⁵ of this procedure using theoretical data, a three-dimensional adsorbate atom structure has been successfully recovered. A further, very recent extension of the idea, backed up once again by numerical simulations, is the analysis of diffuse low-energy electron diffraction (DLEED) intensity distributions⁶. The DLEED method is also applied to adsorbate atoms present at fairly low concentration on random surface sites of an underlying crystal. Hitherto, structural information about the adsorbate atom surroundings has been available only by matching the diffuse background intensity in the diffraction pattern to the results of sophisticated computations of multiple scattering in a whole series of candidate structures.

If these numerical simulations are valid, holographic reconstruction eliminates the need for such elaborate computations and offers a comparatively simple and direct route to surface-structure determination. Further theoretical analysis would be useful to justify, for instance, the claim that multiple scattering effects do not disrupt the reconstruction and to clarify the application of Kirchoff theory to a region that is not completely enclosed by a boundary. More seriously, of course, we

await a rigorous experimental test. The technique should be quite robust to the influence of vibration and stray fields, but the effects of multiple scattering and inelastic or phonon scattering are less clear. Suitable experimental data is readily to hand, for example the striking so-called ADAM (angular-distribution Auger microscopy) images¹⁰ which have been somewhat naively interpreted as silhouettes of surface atoms, with the remarkable assertion that diffraction effects are not relevant. It may be fair to assume, as is apparently done here, that reconstruction with a converging spherical wave is little different from reconstruction with a wave expanding from a small source at the centre of the hemispherical hologram. 'A ball bearing model of the atomic positions' is however a most unconvincing answer to the interesting question of what object characteristic is recovered when an electron hologram is illuminated with light. The high-resolution microscopists have

already found that a naive interpretation of electron holograms could be erroneous. We must hope that the exciting extension of holography to low-energy electron diffraction does not carry with it the same danger. □

Archie Howie is in the Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK.

1. Gabor, D. *Nature* **161**, 777–778 (1948).
2. Tonomura, A. *Rev. mod. Phys.* **59**, 639–669 (1987).
3. Lichte, H. *Ultramicroscopy* **20**, 293–304 (1986).
4. Szöke, A. in *Short Wavelength Coherent Radiation: Generation and Applications* (eds Attwood, D.T. & Boker, J.) 361–367 (Am. Inst. of Physics, New York, 1986).
5. Barton, J.J. *Phys. Rev. Lett.* **61**, 1365–1359 (1988).
6. Saldin, D.K. & de Andres, P.L. *Phys. Rev. Lett.* **64**, 1270–1273 (1990).
7. Frabboni, S., Matteucci, G. & Pozzi, G. *Ultramicroscopy* **23**, 29–38 (1987).
8. Banzhof, H., Hermann, K.H. & Lichte, H. in *EUREM 88* (eds Goodhew, P.J. & Dickinson, H.G.) **1**, 263–264 (Inst. of Physics, Bristol, 1988).
9. Lin, J.M. & Cowley, J.M. *Ultramicroscopy* **19**, 179–190 (1986).
10. Frank, D.C., Batina, N., Golden, T., Lu, F. & Hubbard, A.T. *Science* **247**, 182–188 (1990).

GEOCHRONOLOGY

Timescales and telltale corals

Minze Stuiver

WHEN investigating past events, the radiocarbon (^{14}C) dating method is used widely and is usually considered the most reliable for determining ages over the past 50,000 years. Because of its heavy use it is important to identify its imperfections. Bard *et al.* on page 405 of this issue¹ compare the ^{14}C and $^{230}\text{Th}/^{234}\text{U}$ ages of Barbados corals and provide a much-needed calibration of the radiocarbon timescale beyond the range of previous calibrations.

The past decade has seen a shift from radioactive decay counting to direct determination by mass spectrometry of minute isotope amounts. Accelerator mass spectrometry now delivers ^{14}C ages with good precision for milligram-sized C samples, and Bard *et al.*'s refinement of standard mass spectrometric techniques gives coral ^{230}Th and ^{234}U determinations an order of magnitude more precise than those obtained by decay counting. With such improved technology, the assumptions underlying each dating method become increasingly important.

The incorporation of uranium in coral as it forms, and the subsequent radioactive decay of ^{238}U and ^{234}U to ^{230}Th allows uranium–thorium ages to be determined. The $^{230}\text{Th}/^{234}\text{U}$ and $^{238}\text{U}/^{234}\text{U}$ ratios have to be measured, with the initial amount of ^{230}Th calculated from the measured abundance of ^{232}Th . The corals are assumed to represent a closed system for U and Th exchange. This assumption is crucial as U and Th are notoriously mobile in sediments. The absence of variation in the calculated initial $^{238}\text{U}/^{234}\text{U}$ ratios of the

Barbados corals agrees with these closed-system assumptions, and so Bard *et al.* conclude that the U/Th systematics could not have been strongly altered after deposition. Also consistent with a closed system is the similarity of the $^{230}\text{Th}/^{234}\text{U}$ to the calibrated ^{14}C ages of the past 10,000 years. The absence of any indications of exchange is, however, not necessarily proof of a closed system but it seems that post-depositional alteration is unlikely.

Determining a conventional ^{14}C age in the laboratory involves measuring the ^{14}C activity remaining in the sample, normalizing this activity to a fixed $\delta^{13}\text{C}$ value to eliminate the influence of isotope fractionation, and calculating the ratio R of normalized sample activity to the activity of an NBS standard that represents atmospheric ^{14}C activity of the late nineteenth century. The ratio R is then converted to a conventional ^{14}C age using the radioactive decay law. The crux of the matter is, of course, that the atmospheric ^{14}C activity of the late nineteenth century does not necessarily represent the initial ^{14}C activity of samples grown at different times and in diverse reservoirs differing in ^{14}C activity from that of the atmosphere. For the latter effect an estimate of the ^{14}C age difference with the atmosphere (reservoir correction) has to be made; whereas for the former, information from other dating methods is needed.

The 'fine tuning' or calibration of ^{14}C ages has, so far, been achieved using wood dated by dendrochronology (tree-ring counting). Existing wood chronologies

are derived from environments as diverse as the slopes of the White Mountains, Nevada (bristlecone pine), the peatbogs of Ireland (oak) and the alluvial sands of the major rivers of Germany (oak). The longest continuous chronology goes back to 7938 BC (ref. 2).

High-precision radiocarbon dating of wood from these chronologies allows radiocarbon ages to be plotted against AD/BC calibrated (cal) ages. These plots³ are used to convert the radiocarbon age to the AD/BC timescale. Although computerized versions of the calibration process are available³ the process is not ideal because in many instances several solutions are possible for a single ^{14}C age (maximum uncertainty about 200 cal years)⁴.

The ^{14}C age anomalies encountered for the oldest wood samples are fairly large; that is, wood of nearly 10,000 cal yr BP (before present, with 0 BP = AD 1950) has a conventional radiocarbon age of only about 9,000 ^{14}C years. Part of this discrepancy arises from a known 3 per cent shortening of the ^{14}C timescale because the half-life used, agreed upon internationally 40 years ago, is 3 per cent too small, but most of it is caused by an increase in atmospheric ^{14}C activity over time. Do these age anomalies increase further back in time? The $^{230}\text{Th}/^{234}\text{U}$ ages of the Barbados corals show unequivocally that before 9,000 yr BP, the ^{14}C ages are systematically younger than the U–Th ages, with differences of about 3,800 years at 20,000 yr BP.

Although the measuring errors in $^{230}\text{Th}/^{234}\text{U}$ and ^{14}C coral dating are of comparable magnitude, they differ in their areas of applicability because corals are found only in a portion of the marine domain whereas ^{14}C samples are encountered in all carbon reservoirs. To apply the ^{14}C age correction found for corals to the ^{14}C ages of samples in equilibrium with the atmosphere requires that the time history of the offset ^{14}C ages of the atmospheric and oceanic mixed-layer reservoir be known. This offset, about 400 ^{14}C years for the present-day anthropogenically undisturbed mixed layer of the global ocean, has fluctuated during the Holocene by a few hundred years⁵. The change in offset may have been several hundred years larger before 12,000 cal yr BP because of major changes in ocean deep-water formation related to glacial–interglacial climate change. The coral $^{230}\text{Th}/^{234}\text{U}$ results, together with an estimate of the radiocarbon offset between the reservoirs, provides important information on long-term trends in ►

1. Bard, E. *et al.* *Nature* **345**, 405–410 (1990).
2. Becker, B. & Kromer, B. *Abst. Int. Symp. Archaeometry Heidelberg*, Weber D-6900 Heidelberg, 192 (1990).
3. Stuiver, M. & Kra, R. (eds) *Radiocarbon* **28**, 805–1030 (1986).
4. Stuiver, M. in *Climate and Geo-Sciences* (eds Berger, A., Schneider, S. & Duplessy, J.) Cl. NATO ASI Math. and Phys. Sci. **285**, 39–45 (1989).
5. Stuiver, M., Pearson, G.W. & Braziunas, T.F. *Radiocarbon* **28**, 980–1021 (1986).

atmospheric ^{14}C . However, our insufficient knowledge of palaeo-ocean circulation precludes detailed calculations of the variations of the offset with time and thus the fine detail observed in the atmospheric tree-ring ^{14}C record and the associated calibration of the radiocarbon timescale cannot be matched.

Variations in atmospheric ^{14}C are caused by changes in production rate associated with solar wind modulation, change in the geomagnetic dipole intensity and redistribution of ^{14}C between the carbon reservoirs induced by climate change. Bard *et al.* argue that the changes in atmospheric ^{14}C of up to 40 per cent which they derived from the coral data are

so large that a change in geomagnetic intensity seems to be the only reasonable explanation.

The adjusted ^{14}C timescale shifts the records of global glacial climate, currently set at 18,000 cal yr, back to about 21,000–22,000 cal yr. The putative glacial maximum at 18,000 cal yr is already open to dispute; a move of this sort should stimulate further discussion of the way that the Milankovitch Earth orbital cycles drive climate variations. □

Minze Stuiver is in the Department of Geological Sciences and the Quaternary Research Center, University of Washington, AK-60, Seattle, Washington 98195, USA.

CRYSTALLOGRAPHY

Steps in a growth business

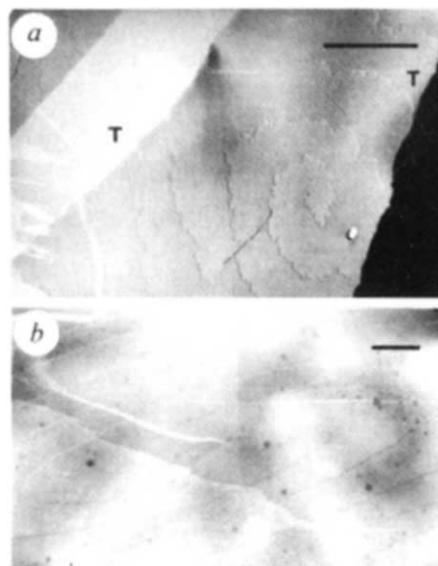
David Dover

THOSE investigating protein structure by X-ray crystallography depend on a reliable supply of sufficiently large, well-ordered crystals on which to work. Improvement of crystallization techniques has been the Cinderella of the business, but has been receiving more attention recently — perhaps because an increasing number of structural studies have stalled owing to crystallization problems. Of the new approaches, the glamour goes to microgravity (in particular) and to robotics, both of which are in the tradition of informed trial-and-error that has been characteristic of the field since its inception¹. More analytical approaches have been tried in a number of laboratories without success. But now that Durbin and Feher² have shown that electron microscopy can reveal much of what is happening, more optimism in the crystal nurseries can be expected.

The advantage of microgravity (growing crystals in satellites in Earth orbit) is that there are no sedimentation or convection currents. If the geometry is chosen correctly, boundaries other than the water–vapour interface can be eliminated too, and the ideal would be to grow crystals in a protein-solution droplet floating freely in space. Although falling short of that ideal, some of the first indications from recent Space Shuttle experiments have been very promising³. These space-grown crystals were larger and better ordered than those grown under normal laboratory conditions. They were brought back to Earth gently, which may be an essential condition for improved results; the dramatic accelerations of unmanned re-entry may itself introduce disorder (or worse). Even if microgravity proves to be efficacious, however, the cost will keep it far beyond the means of the average laboratory.

Back on Earth, more attention is being

given to examining the factors which affect successful nucleation and growth. Robots in the form of automated machines are available both commercially and home made, and will make it much easier to explore the possible conditions for growth. They will be needed no matter how successful analytical tools such as electron microscopy prove — no method for improving the size or internal order of crystals can succeed without initial crystals to examine, and trial-and-error will always be required to provide the first glimmer of crystallization.



a, Step pattern generated by a defect on the face of a lysozyme crystal. The spiral, it is thought, originates at the site where a screw dislocation intersects the surface. The irregular edges (T) are tears in the replica of the surface. *b*, Clearer view of a spiral pattern; steps generated at a defect spiral out and are roughly parallel at large distances. The dark spots and uneven transmission of the replica are due to imperfections in the preparation. Scale bar, 0.5 μm . (From ref. 2.)

Still, many frustrated crystallographers are left wondering why their crystals will not grow large enough or well-ordered enough, or grew yesterday but will not grow today. Durbin and Feher's first results show the 'normal' process of growth very clearly as it is revealed both on the outer surfaces and within the body of the grown crystal. The two basic processes by which material is added to a growing crystal face seem to be the same for proteins as were seen over 20 years ago in paraffins. In both, new molecules are added to a step on the growing face propagating it across the underlying layer until it cannot extend any further, usually at the edge of the face. At high degrees of supersaturation, new steps are produced by nucleation of growing layers on top of existing layers. At low degrees of supersaturation this process is too slow and is outstripped by the growth from step dislocations.

To visualize how such a mechanism works, imagine layers of card, one on top of the other, representing the molecular layers of a perfect crystal, the top one being a crystal face. Cutting through these layers from their edges to their centres results in new edges which extend half way in to the crystal. If the layers are distorted where the cut ends in the centre, so that the cut edge of each layer joins the cut edge of the upper layer adjacent to it, instead of the layer at the same level as before, a step forms across half the crystal face. This step can grow, but as it grows it must spiral around the central dislocation. This is just the spiral growth pattern that Durbin and Feher observe (see figure), and which in other micrographs can be seen propagating through the body of the crystal. Of more significance to those whose crystals won't grow large enough, Durbin and Feher believe that they can also see where steps collide and jam, resulting in cessation of growth.

Many attempts have been made to encourage crystals to grow larger. Etching back the stopped surface and starting again in fresh protein solution has been tried with some success⁴. Again, however, the methods used have been almost all of the trial-and-error variety. Treading in the steps of Durbin and Feher, crystallographers can now settle down to analysis of their own material in the real hope of understanding the events that are increasing the disorder or stopping the growth of their crystals. □

David Dover is in the Department of Biophysics, King's College London, 26–29 Drury Lane, London WC2B 5RL, UK.

1. McPherson, A. *Preparation and Analysis of Protein Crystals* (Wiley, New York, 1982).
2. Durbin, S.D. & Feher, G. *J. molec. Biol.* **212**, 763–774 (1990).
3. DeLucas, L.J. *et al. Science* **246**, 651–654 (1989).
4. Thaller, C. *et al. Meth. Enzymol.* **114**, 132–135 (1985).

Radon and the risks of cancer

Julian Peto

AT the end of April, Denis Henshaw and colleagues¹ published their finding of significant correlations between average indoor radon levels in different countries and the incidence of leukaemia and other cancers in children, and of myeloid leukaemia, kidney cancer and melanoma in adults. Last week, again in *The Lancet*, they announced a similar correlation for prostatic cancer². Significant regional correlations have also been reported between indoor radon and acute myeloid leukaemia in Britain³ and Canada⁴.

Two letters in this week's *Lancet* add fresh grist to the debate. Using data from their atlas of regional leukaemia and lymphoma rates in England and Wales⁵, Ray Cartwright's group describes several significant correlations of cancers — including acute myeloid leukaemia in adults and acute lymphoblastic leukaemia in children — with indoor radon⁶. Scientists from the National Radiological Protection Board (NRPB) however cast doubt on Henshaw's data and conclusions⁷. They observe that both records of cancer incidence and estimates of average radon exposure are of questionable accuracy for some countries, and that the international correlations between adult cancer rates and radon are lower and not statistically significant when the less reliable data are excluded. The correlation for childhood cancers is not reduced in this re-analysis. But the authors point out that the NRPB risk-assessment model for childhood leukaemia predicts that the effect of radon is undetectably small, and argue that in view of this conflicting evidence such correlations constitute only tenuous evidence of cause and effect.

Inconclusive findings

Epidemiological studies on radiation and childhood cancers, particularly leukaemia, have produced a series of intriguing but inconclusive findings in recent years. The discovery in 1983 of a cluster of childhood leukaemias near the nuclear reprocessing plant at Sellafield led first to the report of the advisory group chaired by Douglas Black⁸, and then to the establishment of the Committee on Medical Aspects of Radiation in the Environment (COMARE). The three reports that COMARE has already published confirm that there were unusual clusters of childhood leukaemia near both Sellafield and Dounreay (the only other British reprocessing plant), and that the incidence was also slightly increased in the area surrounding Aldermaston⁹⁻¹⁰. This last observation was supported by an analysis from Richard Doll's group¹¹ of leukaemia mortality below the age of 25 in different

regions of England and Wales, which revealed a small but statistically significant increase in areas close to a nuclear installation.

By this stage the various associations seemed to constitute good evidence of a link, but it was difficult to attribute them to known discharges. Sellafield had emitted much larger amounts of radioactivity than Dounreay, yet the excesses in the local leukaemia rates were similar, and levels of environmental radiation in areas surrounding most other nuclear plants were so low that they seemed unlikely to be the direct cause of any detectable excess. Various other studies provided evidence of unexplained clustering in areas distant from nuclear plants, and the leukaemia death rate below age 25 throughout England and Wales was found to be greater in areas where a high proportion of the population were of high social class¹¹. This could not be accounted for by the difference in leukaemia mortality between social classes, which is minimal in the latest national statistics, and raised the suspicion that something quite different might be involved.

This suspicion was confirmed last year by Leo Kinlen. Kinlen demonstrated¹² that isolated communities sometimes suffer a marked increase in childhood leukaemia after a sudden population influx, and he suggested that an anomalous pattern of childhood exposure to an infectious agent could be the cause. The communities near Sellafield and Dounreay are extreme examples of such population change.

The growing belief that the apparent associations with nuclear installations were more likely to be related to early infections than to radiation was, however, unexpectedly challenged three months ago. Martin Gardner and his colleagues¹³ found that all four of the fathers of children with leukaemia living near Sellafield for whom radiation records could be traced had suffered higher radiation exposure than most other workers at Sellafield. Although the association is based on only four cases, it is highly statistically significant, and is the more worrying because the lifetime exposures of the men involved were within approved international limits. These leukaemias achieve statistical significance by two completely independent criteria: first, the high local incidence, which would not have been significantly elevated if these four cases had not occurred; and second, the high external exposure of their fathers to gamma radiation compared with other employees at Sellafield.

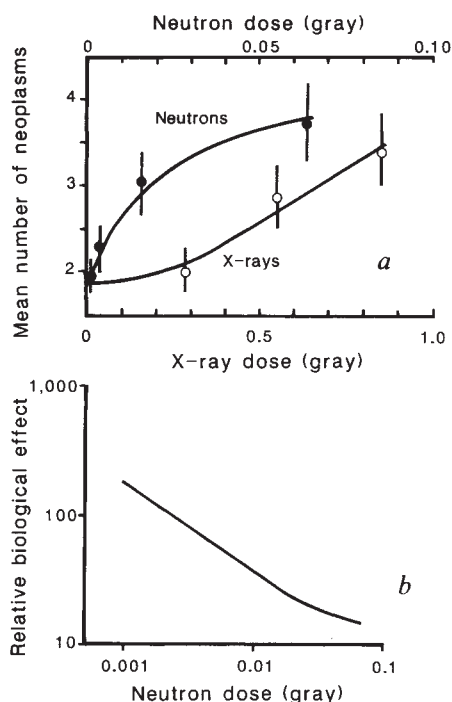
The Gardner study thus seems to

provide quite strong evidence that the children of the most heavily exposed workers at Sellafield suffered an increased risk, although a direct effect of paternal exposure to gamma radiation is only one of several possible explanations. As well as chance, other explanations include exposure to mutagenic chemicals, ingestion of alpha-emitting radionuclides and the possibility that such exposures can activate latent leukaemogenic viruses¹⁴. The most heavily exposed workers' families might also be the most mobile, and hence particularly liable to the effect discovered by Kinlen. Studies are now underway on the children of similarly exposed workers at other plants.

Risk estimates

Henshaw's hypothesis that radon causes a significant proportion of childhood leukaemia and adult myeloid leukaemia (and possibly some other cancers) appears to contradict conventional risk estimates, but such estimates are based on a number of questionable assumptions, particularly for internal high-LET (linear energy transfer) emitters and childhood cancers. The conventional 'quality factor' is 20 for high-LET radiation such as alpha particles or neutrons, which means that their actual energy transfer to tissue must be multiplied by 20 to calculate the biologically equivalent dose of gamma radiation. Henshaw and his colleagues point out that high radon levels are associated with increased external gamma radiation, so any correlation with radon is likely to be due in part to the gamma dose, and they conclude that the most plausible interpretation of their findings is that the quality factor for myeloid leukaemia induction by radon daughters is higher than 20 but less than 180. A value of 20 is certainly too low at low doses, although too high at high doses, in some (although not all) experimental systems, reflecting the very different dose-response curves sometimes exhibited by high-LET and low-LET radiation¹⁵ (see figure, over). The assumption that high- and low-LET radiations can be treated as equivalent by using a single conversion factor is a recognized weakness in conventional risk assessment. In certain respects these exposures appear to act as qualitatively different carcinogens, and their effects should really be modelled separately.

A further complication is the internal distributions of different radionuclides. These distributions have not been well characterized, but elements certainly vary in their tendency to concentrate in particular tissues, and presumably in their effect on the haemopoietic stem cells from which leukaemia develops. In principle, the distribution of radon is easier to model than that of other radionuclides — it is chemically inert and its radioactive daughters are short-lived, so its pathway



a. Dose-response curves for mammary tumour induction in female Sprague-Dawley rats, exhibiting downward curvature for neutrons and upward curvature for X-rays at low doses. (Redrawn from ref. 15.) The dose scales for neutrons and X-rays differ by a factor of 10. b. The authors' estimate of the corresponding relationship between relative biological effect (RBE) and dose shows that the RBE exceeds 100 at 0.001 Gy, and suggests that it may be even higher at lower doses.

in the body is determined largely by solubility.

Predictions of lifetime cancer risk to the general population resulting from very low exposures, such as mesothelioma from asbestos in buildings or leukaemia from average radon levels, are usually so low (about 1 in 100,000 or less) that they cannot in principle be confirmed epidemiologically. Moreover, they are based on dose estimates which for certain types and routes of exposure have not been monitored adequately. A model of carcinogenesis that makes untestable predictions from inadequate measurements is more akin to social than natural science, and an epidemiologist confronted with an apparent contradiction between the predictions of radiobiological theory and the incidence of human cancer is bound to suspect that the human data are more reliable. The extraordinary if tiny local clusters of childhood leukaemia at the only nuclear reprocessing sites in Britain, the weaker but still significant excess around Aldermaston, Gardner's evidence that the children of heavily exposed Sellafield workers are at high risk of leukaemia, and Henshaw's correlations between childhood cancer and radon are examples of such contradictions. Some of these observations may be due to chance, and some may be completely unrelated to radiation, as Kinlen has demonstrated;

but there must remain a substantial probability that some are due to unpredicted carcinogenic effects of radiation. In view of Gardner's findings, one possibility is that childhood cancer is caused by paternal exposure to radon.

At a practical level, Henshaw's findings, even if confirmed, merely reinforce the generally accepted conclusion that indoor exposure to high levels of radon is a considerable environmental hazard. The NRPB estimate that as many as 2,000 lung cancers may be caused by radon each year in Britain⁶ provides sufficient grounds for reducing exposure where practicable, irrespective of any added risk for childhood cancer and the other adult cancers for which Henshaw reports significant correlation with radon exposure. The central question is whether or not these correlations constitute strong evidence of cause and effect, and in the absence of clear epidemiological evidence or a much better quantitative understanding of the mechanism of cancer induction the answer must depend largely on one's prior belief in the causal link. Henshaw and his colleagues have been studying the internal distribution of alpha-emitting radionuclides, and conclude that the dose to the red bone marrow and fetus from absorbed radon and its daughters may be substantial¹.

Any scientist would be excited to discover epidemiological correlations that appear to support a hypothesis that emerged from painstaking laboratory studies, and it would be unreasonable to complain because Henshaw and his colleagues are convinced that the relationships they have observed are real. Subsequent research may show that they are wrong. But they have assembled a *prima facie* case for an unexpected, intriguing and testable hypothesis that must certainly be pursued. □

Julian Peto is in the Section of Epidemiology, Institute of Cancer Research, Belmont, Surrey SM2 5NG, UK.

1. Henshaw, D. L., Eatough, J. P. & Richardson, R. B. *Lancet* **335**, 1008-1012 (1990).
2. Eatough, J. P. & Henshaw, D. L. *Lancet* **335**, 1292 (1990).
3. Lucie, N. P. *Lancet* **II**, 99 (1989).
4. Cartwright, R. A., Alexander, F. E., McKinney, P. A. & Ricketts, T. J. *Leukaemia and Lymphoma: An Atlas of Distribution within Areas of England and Wales 1984-1988* (Leukaemia Research Fund, London 1990).
5. Alexander, F. E., McKinney, P. A. & Cartwright, R. A. *Lancet* 2 June 1990.
6. Butland, B. K., Muirhead, C. R. & Draper, G. J. *Lancet* 2 June 1990.
7. Black, D. *Investigation of the Possible Increased Incidence of Cancer in West Cumbria*, (HMSO, London, 1984).
8. COMARE First Report (HMSO, London, 1986).
9. COMARE Second Report (HMSO, London, 1988).
10. COMARE Third Report (HMSO, London, 1989).
11. Cook-Mozaffari, P. J. et al. *Br. J. Cancer* **59**, 476-485 (1989).
12. Kinlen, L. J. in *Medical Responses to Effects of Ionising Radiation* (eds Crosbie, W. A. & Gittus, J. H.) 272-278 (Elsevier Applied Science, London, 1989).
13. Gardner, M. J. et al. *Br. med. J.* **300**, 423-429 & 429-434 (1990).
14. Evans, H. J. *Nature* **345**, 16-17 (1990).
15. Shellabarger, C. J., Chmelevsky, D. & Kellner, A. M. *J. natn. Cancer Inst.* **64**, 821-833 (1980).

Just keep rolling!

MANY urban roads are permanently choked by parked cars, and ruthless methods are often deployed to discourage them. Daedalus now proposes a method which is some sort of ultimate in efficiency and ruthlessness. He recalls those intriguing liquids with nonlinear viscosity, like silicone putty, starch suspensions and mixtures of sharp sand and water ('quicksand'). Under slow steady pressure they flow easily; but rapid impact locks them up into a momentary solid. A suspension of sand in tar is a common road surface, and DREADCO's chemists are now seeking formulations or additives to endow such a mixture with extremely nonlinear viscosity.

A road of DREADCO's 'Quicktarmac' will not inconvenience travelling vehicles. Under the rapid impact of their fast-moving wheels it will act as a stable and reliable solid. But a motorist trying to park on it will find himself in deep trouble. It will rapidly yield to the steady pressure of his static tyres. His car will sink in, and will soon be firmly stuck. Ideally, it should sink completely out of sight, punishing the errant motorist by total loss of his vehicle and freeing that section of road for legitimate traffic again.

But such a great depth of Quicktarmac would be prohibitively expensive. A thin layer of the material on a conventional concrete base will be almost as good. A parked vehicle will merely sink in up to its axles. Thus entrapped, it can never acquire enough velocity to solidify the Quicktarmac and climb onto the surface again; it can only plough slowly and stickily through the material until it reaches the limit of the no-parking zone. An official release vehicle could perhaps free it by violent impact from behind, giving it sufficient instantaneous velocity to ride up again onto the road surface; but most offenders would rather crawl away, crimson with fury and embarrassment, without such aid.

This elegant principle can clearly be extended. Modified formulations of Quicktarmac could discourage kerb-crawling motorists or enforce minimum-speed regulations on link highways. Pedestrians are also strongly motivated not to linger dangerously on the road, but to cross briskly and purposefully.

Even more useful, Quicktarmac is self-repairing. Scars and potholes will fill themselves in by slow spontaneous flow, while trash and rubble sink out of sight. By the same token, however, traditional road camber would flatten out, and a road laid on a gradient would slowly flow downhill. While one might design such a road as a series of shallow steps separated by cross-ribs, Daedalus acknowledges that his new technology is best kept on the level.

David Jones

What use is "euxeny"?

SIR—Rasmont is not the first person to feel uncomfortable about using the term 'altruism' in the context of animal behaviour, even if he is the first to suggest the neologism "euxeny" as an alternative¹. The problem with his suggestion is that many will interpret his term as meaning that biological insights into the evolution of cooperation are simply not relevant to 'real', human altruism.

In practice, the problem hinges on the definition of the latter, not the former. In animal behaviour we have every justification for ignoring motives and defining altruism purely in terms of its objective consequences as any act which enhances the reproductive success of the recipient at a cost to that of the altruist. There are good reasons for thinking that there are three—and almost certainly only three—types of altruistic behaviour: kin altruism, where the sacrifice of the actor benefits copies of the gene for altruism in the recipient to an extent greater than the cost to itself²; reciprocal altruism, where the actor's sacrifice for the benefit of the recipient is reciprocated by the recipient³; and induced altruism, where an actor performs an altruistic act through error, deceit or coercion^{4,5}. Examples of the latter would be birds who raise the offspring of cuckoos, believing them to be their own. On the objective definition of altruism above, they must be seen as acting altruistically because they have incurred a considerable cost to their own reproductive success and enhanced that of the cuckoo.

As Rasmont implies, what makes 'real', human altruism controversial is the belief that it is not motivated by any kind of self-interest. Of course, neither is it in the case of the apparently odd, residual category of

induced altruism. The birds who raise cuckoos have absolutely no self-interest in doing so. Nevertheless, they have been deceived, and so we seem to have to conclude that it is our consciousness of acting without self-interest which defines 'true', human altruism.

But the fact that politicians routinely claim to act in the public interest, rather than in their own self-interest, is not a reason why voters should be expected to believe them. On the contrary, Anna Freud's classic psychoanalytic study of the motives involved in human altruism concluded that "It remains an open question whether there is such a thing as a genuinely altruistic relation to one's fellow men, in which the gratification of one's own instinct plays no part at all . . .". To separate "euxeny" from human altruism, as proposed by Rasmont, would really make it too easy for the politicians, and my guess is that if behavioural scientists were not human beings themselves they would have long since realized that attempts to defend subjective definitions of human altruism are completely futile. Perhaps this is because there is a little of the politician in all of us.

CHRISTOPHER BADCOCK

*Department of Sociology,
The London School of Economics,
Houghton Street,
Aldwych,
London WC2A 2AE, UK*

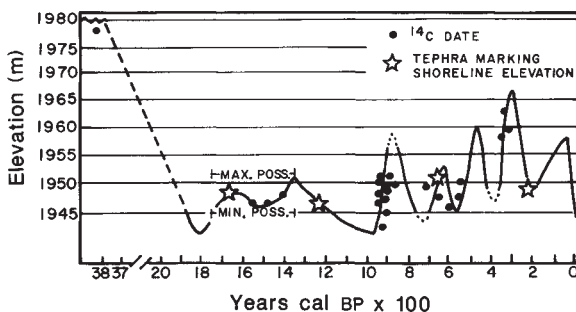
1. Rasmont, R. *Nature* **344**, 498 (1990).
2. Hamilton, W. D. J. *theor. Biol.* **7**, 1–52 (1964).
3. Trivers, R. Q. *Rev. Biol.* **46**(4), 35–57 (1971).
4. Trivers, R. *Social Evolution* (Benjamin/Cummings, Menlo Park, California, 1985).
5. Badcock, C. *The Problem of Altruism* (Blackwell, Oxford, 1986).
6. Freud, A. *The Ego and the Mechanisms of Defence* (Hogarth and the Institute of Psychoanalysis, London, 1979).

Past climate at Mono Lake

SIR—Enzel *et al.* have reported¹ that some time between 4,000 and 3,840 calendar years BP and some time between 522 and 313 cal BP (all dates are radiocarbon dates calibrated by dendrochronology), a substantial body of water covered the now-desiccated Silver Lake Playa in the Mojave Desert. Mono Lake, 380 km to the north, is by contrast a perennial body of saline-alkaline water which can provide a palaeoclimatic record covering a wide range of conditions, from the wettest to the driest².

The artificial recession of Mono Lake in recent decades has made accessible previously drowned evidence of past lake-level fluctuations. At least four generations of

tree and shrub stumps that grew during former low stands lie rooted on newly exposed shore lands. Dozens of research pits have been dug at locations that would, under natural conditions, have been under 10 m of water. Three airfall tephra from nearby volcanoes and a float-line



Late Holocene fluctuation curve for Mono Lake.

of pumice blocks from a sub-lacustrine volcanic eruption have made it possible to reconstruct lake margins at the time of eruptions.

Most importantly, the drop in lake level has forced the three largest tributary streams to incise their deltas to depths of up to 10 metres. The walls of these cuts, each more than 1,000 metres long, show alternating sequences of transgressive and regressive deltaic sediments replete with plant detritus and interspersed with tephra of known age. Individual transgressive and regressive units can often be traced directly to a strand line that marks their terminal elevation. This evidence, combined with twenty-six radiocarbon dates, provides a direct and high-resolution record of the late Holocene fluctuations of the level of Mono Lake³.

In addition to its prominent peaks at approximately 3,770 and approximately 300 cal BP, which appear to correspond to the hydric episodes at Silver Lake Playa, in the late Holocene, Mono Lake underwent five other decades-long transgressions. Analyses using Vorster's hydrological model⁴ indicate that the effective inflow to the Mono Lake groundwater basin required to explain the highest of these transgressions was in excess of 134 per cent of the modern (1937–79) average value. Not apparent from the Silver Lake record are prolonged periods when the climate was dryer than it is now, represented at Mono Lake by recessions and low stands, the most extreme of which occurred in response to decades-long effective inflows averaging less than 68 per cent of the modern value.

Corroboration can be found in the Bristlecone pine tree-ring record not only for the Silver Lake stands¹, but also for the late Holocene fluctuations of Mono Lake². Even more intriguing is the apparent correspondence between the Mono curve and the record of DeVries-type solar variations^{4,5}. This suggests that decades to century-scale climatic changes such as those reflected in the fluctuations of Mono Lake and other lakes of the western Great Basin may be related to such solar events as the Maunder, Spörer, Wolf and mediaeval sunspot minima, and the mediaeval and Roman sunspot maxima^{2,3}.

SCOTT STINE

*Lamont-Doherty Geological
Observatory,
Columbia University,
Palisades,
New York 10964, USA*

1. Enzel, Y., Cayan, R.D., Anderson, R.Y. & Wells, S.G. *Nature* **341**, 44–47 (1989).
2. Stine, S. thesis, University of California (1987).
3. Stine, S. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* (in the press).
4. Stuiver, M. & Quay, P. *Science* **207**, 11–19 (1980).
5. Eddy, J. *Climatic Change* **1**, 173–190 (1977).

Hot solution

SIR—Nelson's report (*Nature* 344, 115–116; 1990) that cables and instrumentation in remote locations can be protected by the application of Tabasco sauce and silicone sealant is a reminder that those who tend orchards also have problems with masticatory and/or curious animals, and that there is a product which is much more potent, easier to apply and (if bought in bulk) cheaper than Tabasco. This is Hot Sauce Animal Repellent (Miller Chemical and Fertilizer Corp., Hanover, Pennsylvania), which is a 2.5 per cent solution of capsaicin extracted from capsicum peppers. I have found this to be effective in protecting young almond trees, although I have heard anecdotal reports of deer that have apparently acquired a taste for this spicy fare. But Nelson's addition of silicone sealant seems to me to be an improvement.

DANIEL B. HRDY

*Division of Infectious Diseases,
University of California Davis Medical
Center,
Sacramento, California 95817, USA*

What are sisters good for?

SIR—An important and prevalent error surfaced in a recent Scientific Correspondence¹, in which Conover questioned why female/female pairings among birds are so rare, given that females in such pairs generally produce fertile eggs. Conover correctly pointed out that pairs of unrelated females would have to produce twice as many offspring as male/female pairs to attain the same average reproductive success per female. His mistake was to assert that sister/sister pairs must produce only 33 per cent more young than male/female pairs. This would be true only under very improbable circumstances, and then only for one of the two sisters.

Suppose, for example, that female/female pairs fledge four young (two per female) and male/female pairs fledge three young, on average per breeding season. Thus, if homosexually paired females are unrelated, each gets credit for two young fledged. If the females are sisters, then according to Conover, their average inclusive fitness should equal that of females paired heterosexually (three young, in this example). This is the result one obtains by applying the erroneous "simple weighted sum" definition of inclusive fitness^{2,3} in which individuals are credited with all of their own offspring plus all of their relative's offspring (devalued by relatedness).

Hamilton⁴ defined inclusive fitness as that portion of individual fitness not attributable to the positive or negative influences of others, plus any harms or benefits conferred on others (weighted by rela-

Methods of leprosy prevention

SIR—Bloom does not discuss the use of leprosy vaccines in his commentary article¹ "Vaccines for the Third World". Is this because vaccines are not the right approach to the eradication of leprosy? The simple, practical, alternative measure is to treat the patients with dapsone, kill the bacteria and prevent the spread of the disease. Where this has been done and measurements carried out there has been a dramatic decline in the disease, for example, in northern Zaria state, northern Nigeria, the prevalence of leprosy fell from 46 per 1,000 in 1953 to 1.6 per 1,000 in 1967. In the village of Igabi it fell over the same period from 67 per 1,000 to 2 per 1,000 (ref. 2).

I have previously presented³ evidence for a decline in incidence. Since then, dapsone resistance has increased but has not caused a resurgence of the disease at least in Katsina Province, northern Nigeria, where the prevalence in 1987 was 2 per 1,000 down from 39 per 1,000 in 1951 (ref. 4). The introduction of multi-drug therapy may lessen the frequency of

resistance to dapsone.

There are still areas where leprosy is endemic and where patients do not have access to this therapy. Even in those areas where it is available the decline in incidence is rarely recorded. In my opinion, failure to extend this treatment and ensure proper evaluation while trying to develop an effective vaccine is irresponsible, especially as the presence of the antibody against the human immunodeficiency virus may make people more vulnerable to leprosy².

C. L. CRAWFORD

*Department of Anatomy,
Charing Cross and Westminster
Medical School,
University of London,
London W6 8RF, UK*

BLOOM REPLIES—Although dapsone therapy has made a significant contribution to the treatment of leprosy in the past, WHO has recommended multidrug therapy with three antibiotics (including dapsone) for several reasons. Multidrug therapy is more effective, has thus far blocked the emergence of drug-resistant organisms that is one of the main problems with single drug treatment, prevents relapses after treatment and can reduce the duration of treatment from a lifetime to a few years or perhaps less. Serious efforts are being made by WHO, the International Leprosy Organization and many endemic countries to extend multidrug therapy to as many patients as possible. Just under half of registered patients worldwide are now receiving this type of treatment. Finally, because of the long latency of the infection, transmission of leprosy frequently occurs before treatment can begin.

The point of developing a vaccine is to prevent infection and disease. It was an omission in my commentary not to mention that four candidate anti-leprosy vaccines are currently in clinical trials.

BARRY R. BLOOM

*Department of Microbiology and
Immunology,
Albert Einstein College of Medicine,
Bronx,
New York 10461, USA*

1. Bloom, B.R. *Nature* 342, 115–120 (1989).
2. Crawford, C.L. *Lepr. Rev.* 40, 159–163 (1969).
3. Crawford, C.L. *Nature* 254, 168–170 (1975).
4. Waaldijk, K. *Lepr. Rev.* 60, 59–61 (1989).
5. Meeran, K. *Br. med. J.* 264, 364–365 (1989).

tedness). The appropriate question is whether a female pairing with her sister benefits her sister sufficiently to offset the cost of not pairing with a male, ($b \geq c/r$). If both sisters could produce the same number of offspring by pairing with an unrelated female, then there would be no benefit to pairing with a sister. In fact, each would suffer an inclusive fitness decrement due to the cost her sister would incur by not pairing with a male. Only in extremely asymmetrical relationships would one of the sisters in a homosexual pair get credit for as many genetic replicates as a female in a male/female pair (again, using the above hypothetical numbers of fledglings). Specifically, if a given female could fledge a total of four young with either an unrelated female or her sister, but her sister's only reproductive opportunity were to pair with her, then her inclusive fitness would equal three young if she paired with her sister versus two young if she paired with an unrelated female. Although theoretically possible, the situation is not likely to occur in nature.

GREGORY F. GREYER
ANNALISA M. WEAVER

*Animal Behaviour Group,
Department of Anthropology,
University of California,
Davis,
California 95616, USA*

1. Conover, M.R. *Nature* 342, 624–625 (1989).
2. Grafen, A. *Nature* 298, 425–426 (1982).
3. Grafen, A. in *Behavioural Ecology: An Evolutionary Approach* 2nd edn (eds Krebs, J.R. & Davies, N.B.) 62–84 (Blackwell, Oxford, 1984).
4. Hamilton, W.D. *J. theor. Biol.* 7, 1–52 (1964).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of fewer than 500 words and five references. □

Ice-age dust and sea salt

SIR—Harvey¹ has estimated the climatic effect of increased tropospheric aerosols during the last glacial maximum (LGM). Attributing to the present tropospheric aerosol a globally averaged optical depth, τ , of 0.13 (at 550-nm wavelength) and a 2–3-K surface cooling, Harvey rather more than doubles τ during the LGM and finds a further cooling of 2–3 K.

The ice-core evidence he cites concerns only two aerosol components: continental dust, which is measured by the insoluble particles Al or Ca²⁺; and marine sea salt, which is measured by Na⁺. Together, dust and sea salt form the 'coarse mode', or supermicrometre, fraction of today's atmospheric aerosol, with sources, sinks and optical properties distinct from the fine, or submicrometre, mode.

Because these two modes vary independently in the atmosphere (which the ice-core record itself dramatically confirms²), Harvey's modelling exercise properly constitutes an estimate of the LGM cooling resulting from dust and sea salt only, and should therefore have begun with an estimate of the optical depth due solely to the coarse mode, τ_{coarse} , in the present-day climate. Note that the studies cited for confirmation of his control model were each concerned with the total tropospheric aerosol, not just the coarse mode fraction.

Early investigators (see ref. 3) recognized that the tropospheric aerosol mass is more or less evenly divided between primary particles, mainly dust and sea salt, and secondary particles (those produced *in situ* from gases), mostly sulphates. It was later demonstrated that these primary and secondary components

have very different sizes^{3,5} and, consequently, very different solar extinction efficiencies⁶, the fine-mode sulphates being about six times more efficient per unit mass. Thus, a reasonable estimate, in agreement with limited direct observations, is that the coarse mode causes about 15 per cent of τ . (It will cause an even smaller fraction of aerosol-induced cooling since larger particles scatter sunlight much more into the forward direction.)

Taking this estimate as uncertain by a factor of 2 and assuming that current τ falls somewhere between 0.1 and 0.2 (ref. 7), we can estimate τ_{coarse} to be within the range 0.01 to 0.06. Thus, Harvey's calculation, based on a control model τ of 0.13, should be considered as an extreme upper limit of the climatic forcing due to ice-age dust and sea salt. Because Harvey's estimated LGM cooling of 2–3 K is simply proportional to the value of τ used in his control model, it should be reduced by a factor of at least two and possibly ten.

T. L. ANDERSON

R. J. CHARLSON

Department of Atmospheric
Sciences,
University of Washington,
Seattle,
Washington 98195, USA

1. Harvey, L. D. *Nature* **334**, 333–335 (1988).
2. Legrand, M. R., Delmas, R. J. & Charlson, R. J. *Nature* **334**, 418–420 (1988).
3. Jaenicke, R. *J. aerosol Sci.* **11**, 577–588 (1980).
4. Whitby, K. T. *Atmos. Environ.* **12**, 135–159 (1978).
5. Clarke, A. D., Ahlquist, N. C. & Covert, D. S. *J. geophys. Res.* **92**, 4179–4190 (1987).
6. Willeke, K. & Brockmann, J. E. *Atmos. Environ.* **11**, 995–999 (1977).
7. Toon, O. B. & Pollack, J. B. *J. appl. Meteorol.* **15**, 225–246 (1978).

Chance marsupial relationships

SIR—Thomas *et al.*¹ interpret their evolutionary tree, based on 12S ribosomal RNA gene sequence data for the marsupial wolf (*Thylacinus cynocephalus*) and six other marsupials, as evidence that the thylacine does not lie outside the Australian cluster of marsupial lineages. The strength of this evidence depends on the confidence we can have in their tree topology (their Fig. 2), but a comparison of the cladistic analysis of the observed 12S ribosomal RNA data with similar randomized data sets suggests that their tree is not significantly different in its degree of cladistic structure, from cladograms produced 'by chance alone'.

The length of a cladogram (the smallest number of character state changes implied by the tree of minimum length) is a measure of cladistic structure. It reflects the extent to which co-variation of the characters allows hypotheses of common ancestry, as represented by sister-group

relations in the cladogram, to explain shared character states among the taxa^{2,3}. Because cladistic algorithms will impose some cladistic structure even on data consisting of random numbers, it is informative to ask whether a given data set has more cladistic structure than would be expected by chance.

The original cladistic structure can be compared to the structure revealed by artificial data sets in which each original character is taken in turn and its observed character states are randomly reassigned, without replacement, to the taxa. A Monte Carlo study (repeated computer simulations using random numbers) compares the observed data to many such randomized sets.

The Monte Carlo test for the 12S ribosomal RNA data focuses on the relationship of thylacines to other accepted monophyletic groups of marsupials. The two taxa in the *Dasyuridae* and the two in *Phalangeridae* are each combined into a single family-level taxon, to avoid testing for structure in the data due simply to the monophyly of these groups. Thus, five marsupial taxa are included in the data set. Because the outgroup status of *Phalangeridae* (*Didelphidae*) is well supported⁴, its character states are held constant; it is used as an outgroup taxon in the analysis of the randomized sets.

Characters (ribosomal RNA sites) that are invariant or which differ by a single character state among the five taxa are ignored; these will contribute constant amounts of character change over all data sets, and will not affect the results of the test. The reduced data set therefore has seven phylogenetically informative characters for the five taxa (see table). Cladistic analysis recovers the same topological relationships between the five families as reported by Thomas *et al.*; the tree has nine steps.

One of the randomized data sets from the Monte Carlo study is also shown in the table (data in parentheses and bold type). From a total of 99 randomized sets, a tree of the same length as, or shorter than, that of Thomas *et al.* is found nine times. A rough estimate of the significance is the proportion of all trees (including the original) having nine steps or fewer, that is 0.10.

The results of this Monte Carlo test are surprising. They imply that if the 12S ribosomal RNA nucleotides at each of the seven sites had been randomly assigned to taxa, there would have been a reasonable chance (about 1 in 10) that the cladogram produced would have had as much struc-

CHARACTER STATES AT PHYLOGENETICALLY INFORMATIVE SITES

Taxon	1	2	3	Site 4	5	6	7
<i>Didelphidae</i>	G(G)	C(C)	C(C)	T(T)	A(A)	C(C)	C(C)
<i>Thylacinidae</i>	A(A)	C(C)	T(T)	A(T)	A(T)	T(C)	T(C)
<i>Dasyuridae</i>	G, A(A)	C(A)	T(T)	A(T)	A(T)	T(T)	T(T)
<i>Peramelidae</i>	G(G)	A(A , T)	C(T)	T(A)	T(A)	C(C)	C(T)
<i>Phalangeridae</i>	A(G , A)	A, T(C)	T(C)	T(A)	T(A)	C(T)	C(C)

Nucleotides at each of the seven informative sites (results of a Monte Carlo simulation in parentheses and bold type). For two of the sites in each array, one of the taxa is assigned two different nucleotides. The necessary extra character state change within the taxon is ignored in calculating tree length. The taxon assigned 'A, T' at site 2 is parsimoniously interpreted as having ancestral state 'A', and that assigned 'G, A' at site 1 as having that ancestral state implying the smallest number of character state changes for a given tree.

ture as that in Thomas *et al.*'s cladogram. I conclude that the 12S ribosomal RNA data constitute insufficient evidence to determine the phylogenetic relationships among these taxa, and in particular that the data do not exclude the possibility that the thylacine is an outgroup to the Australian lineages.

Our more formal presentation of the Monte Carlo test and its application to other taxonomic data sets⁵ supports the recent view that, for both animals⁶ and plants⁷, the degree of character convergence and reversal found in molecular data sets matches that of morphological data. This property, together with the large number of phylogenetically uninformative characters found in most molecular data sets, suggests that the apparent wealth of molecular information for phylogenetics is frequently illusory.

DANIEL P. FAITH

Division of Wildlife and Ecology,
CSIRO,
PO Box 84, Lyneham,
ACT 2602,
Australia

THOMAS *ET AL.* REPLY—Faith's randomization test, which is similar to one recently published by Archie⁸, demonstrates the absence of any significant departure from randomness for the 12S ribosomal RNA gene sequences taken as a whole. But for testing if particular groups of taxa in a tree are monophyletic or not, the bootstrapping approach is more powerful⁹. For the particular question addressed in our paper¹, the latter test gives significant support for the thylacine-dasyurid clade. Other portions of the tree, such as the position of the bandicoot *Echymipera*, are not well supported by the 12S ribosomal RNA gene data when tested by bootstrapping.

Furthermore, confidence in the existence of the thylacine-dasyurid clade is enhanced by the quantitative immunological results obtained with albumin⁹ and by the finding that the ratio of transitions to transversions observed in the cytochrome *b* gene is higher for comparisons of Australian marsupial carnivores with each other and the thylacine than for all comparisons of these marsupials with others¹. In an attempt to achieve even higher confidence, we intend to obtain more sequences from the thylacine and extant marsupials.

As for Faith's view that molecular traits are not superior to morphological ones for building reliable genealogical trees, we refer readers to the case of the orang utan's position among hominoid primates. Despite more than a century of intensive morphological analysis, no statistically significant support emerged for any of the three following trees, in which the orang utan is less akin than humans are to Afri-

can apes; the orang utan is more akin to humans than African apes are; or the orang utan is more akin than humans are to African apes^{10,11}. Yet studies of DNA reveal 121 traits supporting tree A, 8 supporting tree B and 10 supporting tree C — a highly significant result¹¹. A possible reason for Faith's pessimistic view is that the cases to which he refers are for small proteins whose resolving power is inherently low when used to probe deep branches in genealogical trees.

RICHARD H. THOMAS

Department of Genetics,
University of Nottingham
Nottingham NG7 2UH, UK

SVANTE PÄÄBO

ALLAN C. WILSON

Division of Biochemistry and
Molecular Biology,
University of California,
Berkeley,
California 94720, USA

1. Thomas, R.H., Schaffner, W., Wilson, A.C. & Paabo, S. *Nature* **340**, 465–467 (1989).
2. Farris, J.S. *Syst. Zool.* **28**, 483–519 (1979).
3. Faith, D.P. *Cladistics* **5**, 235–258 (1989).
4. Clemens, W.A., Richardson, B.J. & Baverstock, P.R. in *Fauna of Australia Mammalia* (eds Walton, D.W. & Richardson, B.J.) 527–548 (AGPS, Canberra, 1989).
5. Faith, D.P. & Cranston, P.S. *Cladistics* (in the press).
6. Wyss, A.R., Novacek, M.J. & McKenna, M.C. *Molec. Biol. Evol.* **4**, 99–116 (1987).
7. Syvanen, M., Hartman, H. & Stevens, P.F. *J. molec. Evol.* **28**, 536–544 (1989).
8. Archie, J.W. *Syst. Zool.* **38**, 239–252 (1989).
9. Lowenstein, J.M., Sarich, V.M. & Richardson, B.J. *Nature* **291**, 409–411 (1981).
10. Wilson, A.C. *Scient. Am.* **253**(4), 164–173 (1985).
11. Wilson, A.C. *et al.* in *The Hierarchy of Life* (eds Fernholm, B. *et al.*) 407–419 (Elsevier, Amsterdam, 1989).

Papain and related proteins

SIR—Recent contributions to Scientific Correspondence^{1–3} have addressed the structural similarity of the 111K antigen of the malaria parasite *Plasmodium falciparum* to cysteine proteinases in the amino-acid sequence region containing Cys 22 and Cys 25 in papain. The fact that Cys 585 of the 111K antigen is analogous to Cys 22 of papain^{2,3} rather than to the catalytic site Cys 25 (ref. 1), for which the analogous residue is Ser 588, prompted Mottram *et al.* to suggest¹ that the 111K antigen might represent an evolutionary intermediate between cysteine and serine proteinases.

In considering the possibilities for variation in the catalytic-site regions of cysteine proteinases and structurally related proteins, our isolation of a protein in a metastable configuration from the latex of *Carica papaya* in which first, the sulphur atom of Cys 25 probably exists as a disulphide with the sulphur atom of either Cys 22 or Cys 63; and second, one or other of these residues contains a free thiol group, may be of interest.

In papain, the side chains of Cys 25 and

His 159 form an interactive system with nucleophilic character in weakly acidic solution modulated by binding interactions and electrostatic field effects^{4,5}. The S^-/Im^+H ion-pair component of this interactive system is probably the catalytic device responsible for the nucleophilic and general acid attack on the scissile bond of a substrate⁶. The side chains of Cys 22 and Cys 63 form one of the three disulphide bonds of the papain molecule, and analogous features are known to exist in other cysteine proteinases.

The reactivity of the thiol group of Cys 25 in active papain towards the two-hydronic-state reactivity probe⁷ 2,2'-dipyridyl disulphide (2-Py-S-S-2-Py) is abnormally high in acidic media and maximal at pH 3.8. The abnormal reactivity derives from the coexistence at pH 3–4 of significant concentrations of (Cys 25)- $S^-/(His\ 159)-Im^+H$ ion pair and the highly electrophilic 2-Py-S-S-2-Py⁺H cation⁸.

In marked contrast, the thiol group of the metastable papaya latex protein ('propapain') exhibits more normal reactivity characteristics typical of a 'bare' thiol group that is not part of a Cys-His interactive system, with maximal reactivity in alkaline media and negligible reactivity in acidic media^{8,9}. Treatment with a low molecular mass mercaptan such as cysteine converts the inactive, thiol-containing, metastable protein into catalytically active papain, with retention of thiol stoichiometry and with marked change in thiol reactivity characteristics to those typical of cysteine proteinase catalytic sites⁹.

The results of kinetic⁸ and labelling⁹ experiments support the view that the metastable protein is a structural isomer of papain containing a different disulphide bond (Cys 22–Cys 25 or Cys 25–Cys 63), and that its conversion to papain involves thiol-promoted thiol-disulphide interchange to produce the free thiol group of Cys 25, which forms the catalytic-site ion pair with His 159.

KEITH BROCKLEHURST

Department of Biochemistry,
St Bartholomew's Hospital Medical
College,
University of London,
Charterhouse Square,
London EC1M 6BQ, UK

1. Higgins, D.G., McConnell, D.J. & Sharp, P.M. *Nature* **340**, 604 (1989).
2. Eakin, A.E., Higaki, J.N., McKerrow, J.H. & Craik, C.S. *Nature* **342**, 132 (1989).
3. Mottram, J.C., Coombs, G.H. & North, M.J. *Nature* **342**, 132 (1989).
4. Brocklehurst, K. *et al.* *Biochem. J.* **256**, 543–555 (1988).
5. Kowlessur, D. *et al.* *Biochem. J.* **259**, 443–452 (1989).
6. Brocklehurst, K., Willenbrock, F. & Salih, E. in *Hydrolytic Enzymes* (eds Neurberger, A. & Brocklehurst, K.) 39–158 (Elsevier, Amsterdam, 1987).
7. Brocklehurst, K. *Meth. Enzym.* **87C**, 427–469 (1982).
8. Brocklehurst, K. & Kierstan, M.P.J. *Nature new Biol.* **242**, 167–170 (1973).
9. Kierstan, M.P.J., Baines, B.S. & Brocklehurst, K. *Biochem. Soc. Trans.* **10**, 172–173 (1982).

A thousand and one Eves

J. S. Jones

The Search for Eve. By Michael H. Brown. *Harper & Row: 1990. Pp.357. £16.95.*

MITOCHONDRIAL DNA, like Jewishness, is inherited down the female line. As we each have one mother, two grandmothers and so on, then we all must share a remote female ancestor in common. What is true for Jews is also true for Joneses: surnames travel down the generations through males and — except for names which originate more than once — everyone with a particular name can trace it to a common male forebearer. The ur-Jones lived a few hundred years ago; the Ultimate Jewish Mother some centuries before that.

Brown's book is a journalist's account of the search for the mother of us all; the mitochondrion which lay at the root of the human family tree and within the cells of a woman who inevitably became known as "Eve". It is a racy tale about the human race and about the quarrels of those who study it. As usual, the science is more interesting than the scientists. Does it really matter that one proponent has "piercing gelid eyes" while another is a "demure earth mother"? Admittedly, though, there are few fields which can boast papers written from prison where the author is serving time for poisoning the judge

who gave him a drugs sentence. Palaeo-anthropologists seem to make up for a lack of fossils with an excess of fury, and this must now be the only science in which it is still possible to become famous just by having an opinion. As one cynic says, in human palaeontology the consensus depends on who shouts loudest.

The genetics of human ancestry has also led to some raised voices, with plenty of accusations of incompetence and fraud. Although there are some worrying exceptions it does seem that for most genes, including those of the mitochondrion, the ancestral populations are in Africa (or, possibly, in the Middle East). From there, say the geneticists, the first modern humans spread through the world, driving the Neanderthals to extinction without mating with them. This is heresy to many students of the fossil record, who see what appear to be intermediate forms in other parts of the globe. They mock the idea of 'killer humans', spreading like killer bees

and wiping out the "klutzing lame-brain Neanderthals" — to quote one of the many clumsy phrases in this book — as they went.

This mockery is misplaced. First, it is hard to see how pre-human populations

the fact of evolution rather than for how it happened. But although genetics can tell us a lot about the patterns of relatedness of living populations, it can do little more than speculate about the time and place when they split. What comes out of a genetic model often depends on what assumptions go in: to take an unusually rigorous view of molecular evolution — that of the creationist Duane Gish — the fact that man and watermelon are both 97 per cent water does not necessarily mean that they have a recent ancestor in common.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Pick of the bunch — Eve, the mysterious first lady.

thousands of miles apart could independently evolve into the same new species. In addition, invaders — rabbits, grasses, fruitflies or squirrels — can spread at an astonishing rate. They often displace their relatives without interbreeding; a process which, ironically enough, is manifest in the patterns of mitochondrial genotypes in the killer bees introduced into Brazil 30 years ago. There is nothing intrinsically implausible about the 'out of Africa' view of human evolution.

The fossil record can at least tell us exactly who was where and when, but its incompleteness means that it is very likely that no fossil hominid yet found is on the direct line of descent to modern humans. As a molecular biologist quoted in this book says, we know that modern genes came from our ancestors, but we can never be sure that any fossil left a descendant. Fossils are, above all, evidence for

Nothing has caused more argument than the date when various lineages

parted from each other, or when "Eve" might have flourished. New fossil discoveries mean that man has been getting younger every year, and there are now no major disagreements about the date of the separation of the human and the ape family trees. The date of our mitochondrial great-grandmother is much more contentious. The speed at which her descendants diverged from each other depends on two things: the rate of mutation — which can be measured — and the size of the evolving populations, which cannot. Any estimate as to how many early humans there were and hence the date of the mitochondrial mother must be a wild guess. Indeed, the latest estimates of Eve's birth date are between fifty and five hundred thousand years ago, estimates so wide as to be almost meaningless.

Our ancestor's biblical nickname does not, of course, mean that she was then the only female. There were thousands of others, but the accidents of family life,

with some mothers having no daughters and emigrants taking with them only a few of the genes of their homeland when they left, mean that most mitochondrial lineages have been lost. Unfortunately, and in spite of all the publicity about the discovery of the Earth mother, Eve was a statistical artefact.

However, her biography makes a good story. The latest idea is that the spread of one mother's genes accompanied the origin of language; a theory reported here as "scientists claim woman had the first word". This and new work on the Y chromosome means that perhaps we will soon see a book reporting not only where Eve met Adam but what they said to each other. □

J. S. Jones is in the Department of Genetics and Biometry, University College London, London NW1 2HE, UK.

Mirror images

P. V. E. McClintock

Reflections on Liquid Helium. By E. L. Andronikashvili. *AIP*: 1990. Pp. 317. £45, \$60.

THE non-appearance of invited Soviet speakers at international conferences — an all-too-familiar feature of the international scientific scene during the several decades of the cold war — and the resultant relative isolation of Soviet science inevitably led to inefficiencies, infelicities and lost opportunities. One thinks, for example, of the numerous cases of wasteful repetition of work already completed and published on the other side of the iron curtain, of delayed recognition for important achievements, and incorrect attributions of priority in discovery.

Much of the excellent Soviet science of this period has consequently been shrouded in a certain mystery, for those in the West, with many of the chief personae becoming familiar in name only, and years afterwards. This is especially true in the case of the (enormous) Soviet contributions to low-temperature physics, and in particular to the understanding of superfluid helium. The name (but not the person) of Andronikashvili is well known to every low-temperature physicist on account of his famous experiment in which he measured directly the relative proportions of the two components of liquid helium below its superfluid transition temperature at 2.17 kelvin: the normal component (which behaves like an ordinary liquid) and the superfluid component (which displays a variety of exotic properties such as inviscid, frictionless flow).

What Elevter Luarsabovich Androni-

kashvili achieved by managing to complete his *Reflections on Liquid Helium* before his death last year has been, in effect, to "take the lid off" half a century of Soviet low-temperature physics, revealing not only the circumstances of the particular experiment which made him famous, but also the personalities, relationships, arguments, intrigues, motivations and the many difficulties involved in practicing experimental scientific research at the highest level within a closed society. In content, the book represents a kind of scientific autobiography but, as Russell Donnelly points out in his foreword, it is written much more in the style of a novel.

The story opens with Andronikashvili at the age of 29, already established in a chair at the University of Tbilisi in Georgia, but in great trepidation agonizing about whether or not to dare to apply for permission to spend a year in Kapitsa's famous Institute for Physical Problems in Moscow. He was eventually persuaded by his sister-in-law Viva to do so, and the decision changed his life. It also, to a significant extent, modified the future of low-temperature physics. The tale continues with vivid descriptions of Andronikashvili's time in the institute, his enforced return to Tbilisi, and his subsequent (successful) fight to escape from his administrative and political responsibilities for enough of the time to maintain contact with the world of research. Such a succinct summary of the book must make it appear tame but, in reality, it is anything but that.

Most of Andronikashvili's *Reflections* is devoted to discussions of his scientific and social relationships with colleagues, rivals, friends and visitors, and with his hosts in the numerous departments and institutes in the West that he was eventually able to

visit. The story is very much a human one, and is told in large part through quotations of direct speech. The reader may wonder about the accuracy with which conversations are being remembered, in some cases, almost 50 years after the event; but such reservations seem petty in the face of the liveliness, vigour and intrinsic interest of the narrative itself. It provides a fascinating account, from the viewpoint of a participant, of how discoveries are really made — including the errors, arguments, misapprehensions and false starts — as opposed to the tidy *post hoc* rationalization that is usually all that is embodied in a scientific paper. The book will thus be of interest to historians and philosophers of science, as well as to low-temperature physicists.

There are scores of entertaining anecdotes, and there are critical and often highly amusing descriptions of the (literally) hundreds of people whom Andronikashvili met or worked with in the course of his scientific career. The latter include a galaxy of well-known names, such as Kapitsa, Landau, Bogolyubov, Peshkov, Khalatnikov, Frenkel, Shal'nikov, Bohr, Peierls, Allen, Feynman, Bardeen, Abraham, Hall and Vinen — to name only a tiny sample. Andronikashvili's character studies of close associates, for example Kapitsa and Landau, are highly illuminating. In other cases, where his trenchant remarks are based on slighter evidence (such as a single meeting at a conference) they can occasionally seem a little unfair. The persons portrayed are usually highly recognizable, though, and the book will certainly provide a source of considerable amusement to the victims' colleagues and friends.

One of these victims of an Andronikashvili vignette was recently heard to remark good-humouredly: "but he was a pompous little man himself, anyway". Perhaps he was — I never had the privilege (or, possibly, the liability) of meeting him — but he was clearly much more besides. Inevitably, *Reflections* tells us at least as much about its author as about the people he set out to describe. Andronikashvili was clearly a complex and interesting character with coexisting, seemingly contradictory, traits: of impetuous enthusiasm, coupled with meticulous attention to detail; of brusqueness and intolerance, but great personal kindness; of pride and a perhaps overly pronounced sense of hierarchy, mercifully punctured by a well-developed sense of humour. He was also a distinguished scientist who has left a permanent mark on low-temperature physics and on science in Georgia. □

P. V. E. McClintock is in the Physics Department, Lancaster University, Lancaster LA1 4YB, UK.

New Journals

This year *Nature's* annual new journals review supplement will appear in the issue of 11 October. Publishers and learned societies are invited to submit journals for review, taking note of the following criteria:

- Journals which first appeared after June 1988, and which issued at least four separate numbers by the end of April 1990, will be considered for review. The deadline for submission is the end of June.
- Journals covering any aspect of science are eligible, although those dealing with clinical medicine, engineering and pure mathematics are excluded, as are publications of abstracts.
- Frequency of publication must be at least three times a year.
- The main language used must be English. Translation journals in English are eligible.

When submitting journals for review, please send at least four different issues (the first, the most recent and any two others) of each title as soon as possible to: Book Review Editor, *Nature*, 4 Little Essex Street, London WC2R 3LF, UK or 1137 National Press Building, Washington DC 20045, USA.

Right and wrong

Stephen Lock

Beyond Baby M. Edited by Dianne M. Bartels, Reinhard Priester, Dorothy E. Vawter and Arthur L. Caplan. *The Humana Press Inc.*: 1990. Pp. 288. \$29.50 (\$34.50 export).

Philosophical Ethics in Reproductive Medicine. Edited by David R. Bromham, Maureen E. Dalton and Jennifer C. Jackson. *Manchester University Press*: 1990. Pp. 261. £35.

THE public perception of bioethical problems is that they arrive on the scene, fully grown, like Venus from the sea. Surrogate motherhood began to be discussed in the mid 1980s; by May 1988 more than 600 children had been born by this method in the United States. Surrogacy seems unlikely to disappear as a contentious issue, and other current anxieties include the definition of death and the extension of life, embryo research, the allocation of resources (particularly high versus low technology), and the validity of scientific advance for its own sake.

Public concerns make good conference material, but it seems a pity that the organizers of these affairs seem unable to resist recording their deliberations in print. As reviewers never tire of pointing out, the result is inevitably a mixed bag. Some work is original, other findings are not; the papers are of variable quality, even more obviously so when the discussions are included; and some of the content is out of date by the time of publication.

All these comments apply to *Beyond Baby M* and *Philosophical Ethics in Reproductive Medicine*, one from the United States, the other largely British, both based on conferences held two years ago. Sensibly, the US volume omits the discussions and its main thrust is on surrogate motherhood. This practice is not only tedious and costly (averaging \$20,000–35,000 per case) but involves a legal minefield, as shown in the case that gives the book its title. The salient details were well documented in the media: although the Sterns, a childless couple from New York, signed a surrogacy parenting agreement, the biological mother, Mary Beth Whitehead, changed her mind after the birth of baby Melissa. Three court cases resulted in the decision that William Stern and Mary Whitehead are Melissa's legal parents, both with rights and responsibilities

towards her. Given its reputation as a leader in both family law and medical ethics, the forthrightness of the New Jersey Supreme Court was hardly surprising. The real evil was paid surrogacy: "Whatever idealism may have motivated any of the participants the profit motive predominates, permeates, and ultimately govern the transaction" — an aspect recognized early in Britain, where commercial surrogacy was banned in 1985.

The commentaries in this book tease out many of the unappreciated facets of the Baby M case. There is the white/blue collar tension, with the Sterns paying \$60,000–70,000 on legal expenses and private detectives to track the Whiteheads down to their hiding place. There were also family problems — with Mrs Stern having mild multiple sclerosis (not initially revealed to Mrs Whitehead) and Mr

too widely, and the discussions would have been better devoted to printing Iain Chalmers's lecture, published in a journal not readily available, and to providing overviews ("We will leave readers to enjoy tracing these overlaps for themselves", the editors say; well I for one, didn't). To be sure, there are some first-rate accounts: Gordon Dunstan's characteristically wise introduction on the moral status of the embryo, commenting on the remarkable advance in understanding the issues raised by *in vitro* fertilization; Alex Campbell's comparison of withholding neonatal care in the United Kingdom and the United States with case-by-case analyses in the former and an increasing tendency to decision-making by ethics committees in the latter; and Richard Lilford's cogent argument that informed comment is now part of decision theory.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Unhappy families — Mary Beth Whitehead (left), surrogate mother of Baby M. Elizabeth and William Stern (centre and right): potential adoptive mother and biological father.

Whitehead an unsteady employment record and unresolved alcohol problems — not to mention the subsequent divorce of the Whiteheads, with the settlement that he will share in any book or movie royalties. Thus, not surprisingly, a feminist commentator, Barbara Katz Rothman, considers that society is encouraging "production standards" in pregnancy, thinking of the woman herself not as a person but as the container for another, often more valued, person — yet at \$1.57 per hour for her service, the woman is hardly well rewarded.

Disappointingly, in the sole contribution on surrogacy in the British volume, Edgar Page declines to address the feminist perspective, as he does to tease out the argument that if priority is given to the birth mother as against the genetic mother then it is difficult to maintain that mother and father have equal parental status. In fact, with six sections, *Philosophical Ethics in Reproductive Medicine* ranges

Women will run considerable risks for their unborn children. Lilford cites a woman 20 weeks' pregnant found to have cancer of the cervix who was told that immediate treatment sacrificing the pregnancy would carry an 80 per cent chance of complete cure. Asked how delaying treatment until the baby had a 95 per cent chance of survival would reduce her prognosis, he estimated 10 per cent. The woman decided to continue the pregnancy, against her husband's wishes; he urged immediate treatment. (Fortunately, the outcome was successful for both mother and baby, but Lilford argues that even if the mother had died the decision would still have been correct.)

Nowhere is the evanescence of such meetings better illustrated than in the "emergency" discussion of fetal brain cell transplants. The first operation for Parkinson's disease in Britain had taken place in Birmingham only days beforehand and inevitably the proceedings were

tentative, ignoring several issues now known to be crucial (including the unresolved one of efficacy). This is not to criticize the discussants (though were the proceedings worth printing?) but the issue does illustrate the need for some central body to anticipate such developments. The House of Commons in Britain now has no forum for discussing these (though the Lords retains its select committee on science and technology) and it seems extraordinary that Britain has no clear, authoritative source of advice on ethical and policy issues. Several other countries have national bioethics committees — Australia, Denmark, France and Italy — and an outstanding example was the US President's Commission, disappointingly discontinued by President Reagan. It has

been left to private bodies to discuss forming such a committee in the United Kingdom, and at a Nuffield Foundation Conference in April it was agreed that such a committee should involve wide consultation using a discussion document. There is a good case for imitating the evolution of the committee on *in vitro* fertilization in Britain, which has progressed from a voluntary through an interim to an eventual statutory basis. □

Stephen Lock is Editor of the British Medical Journal, BMA House, Tavistock Square, London WC1H 9JR, UK.

■ Also just published is *The New Reproductive Technologies*, edited by M. McNeil, I. Varcoe and S. Yearley. Publisher is Macmillan, price £35 hbk, £12.99 pbk.

Pulling together

Malcolm Irving

Molecular Mechanisms in Muscle Contraction. Edited by J. M. Squire. Macmillan: 1990. Pp. 327. £55.

'MOLECULAR mechanism' has a special meaning in muscle research, because muscle is composed of an array of molecular machines that use the free energy of ATP hydrolysis to produce force or motion. At the heart of the machine are the proteins myosin and actin, the main constituents of the two types of filament which slide past each other when muscles shorten. The subject of this multi-author volume is the structure of actin and myosin filaments and the mechanism of interaction between them. The most popular model for this interaction, the 'rowing' or 'swinging' crossbridge model, proposes that myosin projections transiently bind to the actin filament and rotate about the point of attachment to produce relative filament motion. To quote from the preface, "a major theme of this book is to present the different kinds of structural and mechanical evidence in favour of such a process".

Structural studies dominate the book. An introductory chapter on the basic structural elements of muscle is followed by two more on the structure of the actin monomer and filament. Unfortunately, these were written before the determination of the high-resolution structure of the actin monomer from actin-DNase crystals, which has been discussed at recent meetings and is expected to be published soon. This is likely to have a major impact on the field, and may challenge some of the conclusions and speculations advanced in this book. But the discussion of actin structure here will not be completely superseded, and its main conclusion, that large-scale move-

ments between domains of the actin monomer are possible, could be important in the mechanism of filament sliding. In the swinging crossbridge model, the actin filament is usually considered as a passive track, but this may turn out to be an oversimplification.

There are four chapters devoted to studies of the conformations and movements of the myosin crossbridges in muscle fibres, using X-ray diffraction, electron microscopy, fluorescence and spin probes. It is difficult to compare or link these results, as the techniques report on different structural features and have been applied in different conditions and preparations. (For example, the electron microscopy refers almost exclusively to the flight muscle of the giant waterbug *Lethocerus*, and one of the chapters on X-ray diffraction concentrates on fish muscle.)

Nevertheless, a common conclusion is that the myosin crossbridge may take up more than one conformation while it is attached to the actin filament, as predicted by the swinging crossbridge model. The structures of these attached states, and their relation to the process of force generation, remain poorly defined and it is at least arguable that these studies have

not yet provided firm evidence either for or against the model. An alternative approach is discussed here by Harrington *et al.*, who present a clear and comprehensive account of the evidence for a force-generating mechanism based on a helix-coil transformation in part of the tail of the myosin molecule. Currently the best argument against this hypothesis is the observation that the isolated head fragment of myosin can both move actin filaments and produce the normal force per head, in the absence of the tail. Unfortunately, the chapter seems to have been written before some of these observations were published, so the reader is deprived of the authors' reply to this apparently powerful argument.

Structural studies have shown two attached states of the myosin crossbridge, often called 'weak' and 'strong', by analogy with the binding affinities of the corresponding actin-myosin states in solution. These are discussed in detail in a lucid and thoughtful chapter by Brenner, who also reviews the relationship between solution studies of the actin-myosin ATPase and muscle fibre. The model that emerges is considerably more complicated than the original swinging crossbridge scheme of the 1970s. The contrast between current mechanochemical and structural models for crossbridge function emphasizes how little is yet known of the structures of the transient intermediate states responsible for filament sliding, even at the low resolution of the available structural methods. The low resolution is itself a serious limitation, as might be inferred from the chapter reviewing nuclear magnetic resonance studies of muscle proteins. This chapter discusses the general principles underlying conformational changes in contractile proteins, stressing their dynamic nature and using the role of troponin C in muscle regulation as an example. The corresponding studies of actin-myosin are in a state of relative infancy, but are likely to receive new impetus from knowledge of the crystal structure of actin.

This book does not provide a comprehensive treatment of its subject, and some important approaches are conspicuously absent, notably the more recently developed areas such as molecular genetics and motility studies on isolated actin-myosin. The reference base for the articles extends only to 1988. Each chapter presents a rather personal view of the problem, and these views do not fit together into a neat consensus. On the other hand, each group of authors has the space to develop its ideas in detail and depth, and several of the resulting reviews will, I believe, be of lasting value to those working in the field.

Malcolm Irving is in the Biophysics Department, Kings College London, 26-29 Drury Lane, London WC2B 5RL UK.

New in paperback

■ Two books on dinosaurs have just been issued in paperback. *Digging Dinosaurs*, by John Horner and James Gorman, was reviewed in *Nature* **338**, 552; 1989. Published by Harper and Row, price \$8.95. *Dinosaur Plots* (and other intrigues in natural history), by Leonard Kristalka, is published by Avon Books, price \$8.95. For review see *Nature* **338**, 673; 1989.

■ *Life Among The Scientists*, by Max Charlesworth *et al.*, is subtitled "an anthropological study of an Australian scientific community". Reviewed in *Nature's* Autumn Books issue last year (**342**, 312; 1989), publisher is Oxford University Press, price £8.95.

■ Also just published by Oxford is *The World Within the World* by John Barrow, a wide-ranging study of the laws of nature. Price is £8.95. For review see *Nature* **336**, 288; 1988.

Applications of electron cooling in atomic, nuclear and high-energy physics

Helmut Poth

The introduction of electron cooling has opened up new possibilities in atomic, nuclear and particle physics. Now many newly constructed ion storage rings are making use of this technique.

BEAMS of charged particles are used commonly to probe the fundamental structure of matter and to investigate interactions between elementary particles. As one studies ever smaller scales, the probes must be increasingly energetic and sensitive. Anti-matter projectiles, in particular antiprotons ($\bar{p}$), give access to hitherto inaccessible domains. Because of their opposite charges, one can accelerate counter-rotating beams of particles and anti-particles simultaneously in a single accelerator ring and bring them into collision. The centre-of-mass energies reached in colliding-beam experiments are much higher than those from collisions of projectiles with a target at rest. For instance, one would need to collide antiprotons with energies of 5×10^{12} eV (5,000 GeV) with hydrogen nuclei in a fixed target in order to match the energy of 1,000-GeV colliding $p\bar{p}$ beams. Consequently, the accelerator would have to be 50 times bigger than the collider.

Electron cooling provides a means for focusing such particle beams and reducing their energy spread. These refinements are far more than an experimental nicety: they open up a whole range of investigations that would be otherwise impossible. Electron-positron (e^-e^+) beams are used in collider experiments, but proton-antiproton beams provide higher energies. Antiprotons have to be created artificially. Enough energy is available to produce them in high-energy proton-nucleus collisions, but the yield is low; they have to be accumulated before colliding with protons, if reasonable reaction rates are to be obtained. Moreover, antiprotons so produced are spread out in energy and space, so a single pulse injected into an accumulator ring will fill it up. The pulse must be compressed in order that many pulses may be accumulated in the ring. In a reference frame moving with the beam velocity (the orbiter frame), the energy spread and the divergence correspond to velocity distributions parallel and perpendicular to the beam direction, with corresponding temperatures amounting to millions of degrees. Compression cannot be achieved by standard ion-beam optics, because Liouville's theorem states that the volume occupied by particles in phase (position and momentum) space cannot be reduced by the action of conservative forces (such as those that govern the motion of particles in accelerators and storage rings). According to this principle, the phase-space volume is like a balloon filled with air: squeezing the balloon in one region makes it expand elsewhere. The only way to reduce the volume of the balloon is by cooling the gas. This can be achieved by cooling the material of which the balloon is made, and relying on collisions of the gas with the balloon wall to transmit the cooling to the gas. This is viable as long as the particles do not interact destructively with the wall of the container. But in the case of high-energy beams, wall collisions would result in destruction of the container and projectile, so such a scheme is not feasible. How, then, can the particle beam be cooled? The alternative is to mix it with a cold gas which cools it through direct collisions between the gas particles and the antiprotons. Annihilation has to be avoided, so one needs a gas of particles that interact only electromagnetically with the antiprotons—electrons or positrons would be suitable. Electrons with temperatures of a thousand degrees or so can be obtained easily. This cooling of a beam using a cold electron gas was first

proposed by Budker¹.

Charged particles can also be cooled very efficiently in a different way. Moving charges induce electric fields in antennas when they pass close by, the field being related to their velocity. By feeding the induced signal to capacitor plates through which the charges pass further downstream, a retardation or an acceleration can be imposed on the particles that is dependent on their initial velocity. This method, called stochastic cooling, was invented by van der Meer². The signals are picked up at random from ensembles of charged particles, whence the term 'stochastic'. Stochastic rather than electron cooling is more appropriate for cooling high-energy antiprotons^{3,4}, but electron cooling is used to improve antiproton beams at low energy, and opens up new possibilities in the study of antinucleons and their interaction with ordinary matter. A further important application of electron cooling is as a means of compressing stored beams of light and heavy ions to permit spectroscopic investigation of highly charged ions. Initial studies of this sort used protons⁵⁻⁷, but the method has now also been applied to antiprotons⁸ and light ions⁹, yielding intriguing results.

New possibilities for physics

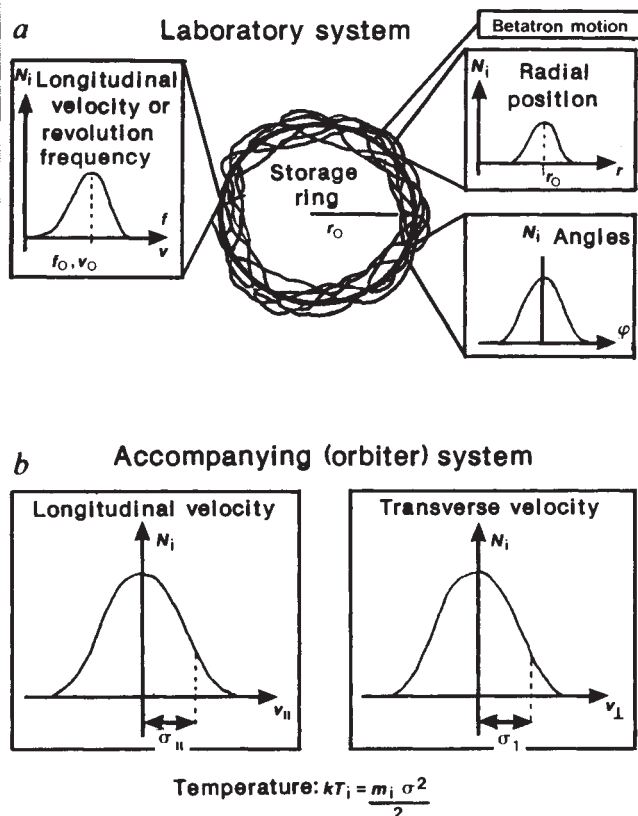
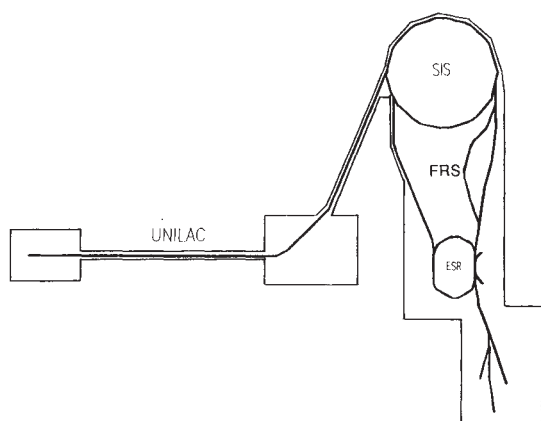
Cooled ion beams are small in diameter and divergence and have a sharp energy distribution, in these respects resembling laser beams. These properties are essential for high spatial and energy resolution, allowing precision experiments with rare species such as antiprotons. Collision and annihilation of antiprotons and protons at low energy, can give information about the interior structure of (anti) nucleons. Such collisions can be initiated with the particles almost at rest, because their opposite electric charges result in a mutual attraction which binds them together in a heavy hydrogen-like atom (a $p\bar{p}$ bound state). On disintegration, their constituent quarks and antiquarks rearrange to form other hadronic particles such as pions. Spectroscopy of these annihilation products yields information on how the quarks and antiquarks are bound together and how they interact with each other. An accurate study of these aspects requires a well defined initial situation (for example, a low energy spread and a well defined interaction angle, which can be obtained by cooling the antiproton beam).

To slow cooled antiprotons it is possible to attach a positron to form antihydrogen. The mass of the antiproton and its gravitational interaction can be measured with high accuracy when antiprotons are confined in electromagnetic traps; electron cooling can help to make capture efficient and to lower the antiprotons temperature. A comparison between the atomic spectroscopy of particle and antiparticle systems enables one to test with great precision the validity of fundamental symmetry relations in particle physics.

For atomic and heavy-ion physics, the perspectives are no less spectacular. Ion sources can produce only partly ionized heavy ions. For instance, to strip uranium atoms of all electrons, they must be accelerated to velocities greater than the velocity of the K electron ($\sim 70\%$ of the speed of light) and passed through a thin foil. The passage through the foil causes the beam to 'blow up'; cooling can compensate for this, resulting in a small, monochromatic beam of fully stripped ions that can

Box 1 The GSI accelerator complex

This device illustrates the production of fully stripped ions and nuclear fragments. Partially stripped ions are accelerated in two stages—first in a linear accelerator (UNILAC) and then in a synchrotron (SIS)—to an energy greater than the stripping threshold of the innermost electron. They may then be either passed through a metal foil to produce fully stripped heavy ions or fired at a target to produce a beam of nuclear fragments (FRS). Either species can be injected into a ring (ESR) for electron cooling. The injected beam has a large spread in longitudinal velocity, and is smeared out in the transverse direction because of 'betatron' oscillations (*a*). In the rotating (orbiter) frame, the particle velocities are distributed around zero (*b*) with a r.m.s. width σ that is related directly to the beam temperature. Electron cooling of the ion beam narrows these distributions and thus reduces the beam size, divergence and momentum spread. The cooled beam can then be accelerated or decelerated as required, or used directly in combination with an internal target in the ring.



be used for the high-precision study of nuclear reactions. Some experiments in atomic physics require slower ions; bare ions decelerated and trapped at thermal energy and confined in a small volume would be ideal for these experiments. The beam blow-up inherent to all means of decelerating beams requires cooling at different energies to maintain the small phase-space volume. Electron cooling is very fast and can compensate for such blow-up when an ion beam repeatedly traverses a thin target inside the ring. This combination of electron cooling and internal target is a new technique that has yet to be fully exploited.

When an energetic ion beam hits a nuclear target, unstable nuclei (nuclear fragments) are produced, and those that have lifetimes of several minutes or more can be collected. As in the case of antiproton production, these initially diffuse ion beams have to be cooled so that many pulses can be accumulated to give high beam intensity. But the ions are slower than antiprotons and can be more easily matched in velocity with electrostatically accelerated electrons. Thus electron cooling is here more appropriate than stochastic cooling because it provides shorter cooling times and smaller beams.

The new GSI (Gesellschaft für Schwerionenforschung) accelerator complex¹⁰ is described in Box 1 as an example to illustrate the production and use of fully stripped ions and nuclear fragments. In this instrument, the high quality of the beam and the combination of electron cooling and an internal target permits nuclei and nuclear reactions to be studied in great detail. There is also the possibility of attaching one electron to the bare nucleus and measuring its binding energy. A single-electron heavy ion is preferred to a multielectron ion for testing the predictions of quantum electrodynamics (QED) because there is no residual electron cloud to complicate the analysis of the energy levels.

In some cases the cooling ring can be used directly as a mass spectrometer; this is particularly useful for measuring the mass of unstable nuclei.

Principles of electron cooling

To effect cooling, the ion beam must be overlapped with cold electrons of the same velocity. The electron energy is therefore lower by a factor m_e/m_i (the ratio of the electron mass to the ion mass, $1/1,836$ for protons). Consequently, ion beams in the MeV energy range can be cooled by keV electrons. Such electron energies may be attained using well established technologies—they are reached in X-ray tubes, for example.

Electron and ion beams are merged in a straight section of the storage ring (see Box 1). In the moving (orbiter) frame the system appears as a hot ion gas mixed with a cold electron gas. Temperature relaxation takes place through Coulomb collisions between electrons and ions. The ions circulate continuously in the storage ring but the electrons are removed at the end of the overlap region, because they can be delivered constantly and so do not need to be recirculated. The cooling effect appears in the laboratory frame as a reduction in the momentum spread (longitudinal cooling)—being slower than electrons, the ions are accelerated slightly by them (faster ions will accelerate the electrons while being themselves decelerated)—and as a reduction of beam size and divergence (transverse cooling), because ions entering obliquely into the parallel electron beam will receive a kick into the forward direction and their transverse velocity component will be reduced.

Because the ions move around the storage ring about a million times per second, there is a frequent repetition of these processes in a short interval, providing cooling times of the order of seconds. Ideally, both beams eventually come into thermal equilibrium, which means that the ion-beam temperature is reduced from several million degrees to about 1,300 K in the transverse direction, and to less than 10 K longitudinally. The transverse temperature is identical to that of the electrons when they leave the thermionic cathode; the longitudinal temperature

is lower, as the longitudinal velocity spread of the electrons, after acceleration, is compressed in the orbiter frame. Thermal equilibrium is reached only under ideal conditions, because the ions scatter from one another frequently when the ion beam becomes very dense, transferring energy from transverse to longitudinal motion; this mechanism acts as a heating process.

We can also consider the electrons as a foil moving with velocity v_0 . The cooling process can then be understood as a cumulative energy loss experienced by the ions due to repeated passages through the foil. The energy loss reduces the ion velocity along the direction of motion until it equals v_0 , with the velocity vector parallel to that of the electrons.

Theory of electron cooling

Considering the cooling process as arising from the friction of ions passing through an electron 'foil', the corresponding energy loss in the moving frame is given simply by the Bohr formula

$$F(v_i) = \frac{dE}{dx} = -4\pi Z_i^2 n_e r_e^2 m_e c^2 L_{\text{Coul}} \frac{c^2}{v_i^2} = -F_0 \frac{c^2}{v_i^2} \quad (1)$$

where n_e is the number density of the electrons, Z_i and v_i are the charge and the velocity of the ion, r_e is the classical electron radius (2.8×10^{-15} m), L_{Coul} is the Coulomb logarithm (a dimensionless constant of order 10 which characterizes the scattering process), E is the ion's energy and F is the retarding force. The equation shows that the energy loss scales inversely with the square of the ion velocity; the quantity F_0 subsumes the preceding cluster of variables. In deriving the frictional or cooling force, which is equivalent to the energy loss, one must take into account the non-zero temperature of the electrons. This requires taking an average over the corresponding electron velocity distribution $f(v_e)$

$$F(v_i) = -F_0 c^2 \int f(v_e) \frac{v_i - v_e}{|v_i - v_e|^3} d^3 v_e \approx -\frac{3}{2} F_0 c^2 \frac{v_i}{v_i^3 + 2\Delta_e^2} \quad (2)$$

This removes the divergence at $v_i = 0$ in equation (1) and leads to a drop in the force for ion velocities below the electron velocity spread Δ_e (Fig. 1). The approximation holds for a spherical Maxwellian electron-velocity distribution with electron temperature $kT_e = \frac{1}{2} m_e \Delta_e^2$. The friction rate (change in velocity per unit time divided by the velocity) is obtained by dividing the force by the ion momentum. The inverse of this quantity is called the cooling time τ . From equation (2)

$$\tau = -\frac{m_i v_i}{F} = \frac{2}{3} \frac{m_i}{F_0 c^2} \frac{v_i^3 + 2\Delta_e^2}{v_i} \quad (3)$$

Thus the cooling time is proportional to v_i^3 (where the velocity is in the orbiter frame) if $v_i \gg \Delta_e$, and is constant for $v_i \ll \Delta_e$. To obtain the cooling time in the laboratory frame, τ_{lab} , we must multiply by the Lorentz factor γ^2 (where $\gamma = 1/\sqrt{1-\beta^2}$ and $\beta = v_i/c$) and by the ratio of ring circumference C to the length of the cooling section, l . If the ion velocity distribution in the orbiter frame is gaussian, v_i can be replaced by the r.m.s. width σ_i to obtain the temperature relaxation time

$$\tau_{\text{lab}} = \frac{C\gamma^2}{l} \tau \approx \frac{4\sqrt{2}}{3} \frac{m_i c}{F_0} \left[\left(\frac{kT_i}{m_i c^2} \right)^{3/2} + 2 \left(\frac{kT_e}{m_e c^2} \right)^{3/2} \right] \frac{C\gamma^2}{l} \quad (4)$$

This formula (except for the factor $C\gamma^2/l$) was presented 50 years ago by Spitzer^{23,24} for temperature relaxation phenomena in plasma physics. It describes the main features of electron cooling, but neglects important details for the cooling of beams with velocity spreads σ_i smaller than the transverse electron velocity spread Δ_e . It represents a conservative estimate, but

describes the basic scaling behaviour of the cooling time, namely its increase as the $\frac{3}{2}$ power of the ion-beam velocity for hot ion beams ($\sigma_i \gg \Delta_e$), its increase with decreasing electron density, and its decrease with ion charge ($\tau_{\text{lab}} \propto m_i/Z_i^2 \approx 1/Z_i$). It can also be seen that at low energies τ_{lab} is nearly constant for constant electron density. At highly relativistic energies, however, τ_{lab} becomes very long because of the γ^2 dependence. For typical parameters ($n_e = 10^8 \text{ cm}^{-3}$, $L_{\text{Coul}} = 10$, $T_e = 1,300 \text{ K}$, $C/l = 25$, $\sigma_i = \Delta_e$, $\gamma = 1$) we obtain $\tau_{\text{lab}} = 1 \text{ s}$ for protons; for fully stripped uranium ions it would be 35 times smaller.

The basic theory of electron cooling^{11,12} was developed using a binary collision model. It has been refined by Sørensen and Bonderup¹⁵, who treated the plasma behaviour of the electrons in more rigorous way. To achieve efficient cooling, a cold electron beam is required, the generation of which is described below.

Electron cooling devices

Essential points in the design of an electron cooling device are the achievement of a low electron-beam temperature, operation under ultrahigh-vacuum conditions, adiabatic acceleration of the electrons to the ion-beam velocity, and 'clean' injection of the electrons into the collector. The electron gun must be designed carefully and a highly stable high-voltage power supply is needed. A longitudinal magnetic field is used to guide the electron beam and to prevent it from blowing up through mutual repulsion of the electrons. In the cooling region, local variations of the magnetic-field direction must be smaller than 0.1 mrad. Powerful vacuum pumps close to the gun and the collector keep the pressure in the cooling region at about 10^{-11} torr or less. At the end of the cooler the electron beam must be decelerated carefully, reducing its power from around several hundred to a few kilowatts, to prevent the collector from being damaged or heated. Good collectors usually capture more than 99.9% of all primary electrons. Figure 2 illustrates a typical electron cooler: it shows the apparatus^{25,26} used for the cooling of slow antiprotons in the Low-Energy Antiproton Ring (LEAR) at CERN.

Electron cooling experiments

The earliest experiments were concerned with measuring the cooling times and frictional forces under various conditions. The results^{5-8,27,28} obtained in different laboratories are remarkably consistent. The time required for reducing the transverse

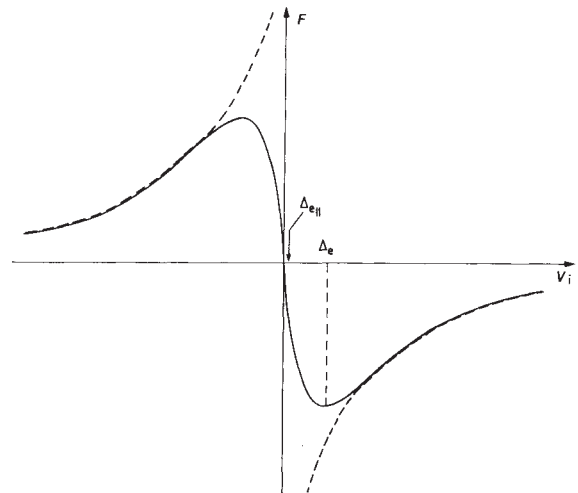


FIG. 1 Cooling force F as a function of the ion velocity in a reference frame moving with the velocity v_0 of the electrons. Dashed line: cooling force for electrons at temperature $T_e = 0 \text{ K}$; solid line: cooling force for electrons at temperature $T_e = \frac{1}{2} m_e \Delta_e^2$. The arrow shows the maximum of the cooling force for a electron velocity distribution compressed in the longitudinal direction to a spread $\Delta_{e\parallel}$.

extension of a beam was measured as a function of the transverse velocity $v_{\perp}$ (Fig. 3a) and was found to be proportional to $v_{\perp}^{-2}$. But equation (3) predicts a $v_{\perp}^{-3}$ dependence for $v_{\perp} \gg \Delta_e$ and no dependence on $v_{\perp}$ for $v_{\perp} \ll \Delta_e$, so this observation that the cooling time decreases continuously with decreasing ion velocity was unexpected. Such '(super)fast' cooling is well described, however, by a more elaborate theory^{12,15} and results from the longitudinal magnetic field (solenoid) in which the electrons move. This field transforms an electron's transverse motion into a spiral about the magnetic field lines (cyclotron motion). Scattering of ions by electrons at collision cross-sections larger than the radius of these cyclotron orbits averages out the transverse electron motion, so that only the longitudinal velocity spread $\Delta_{e\parallel}$ is relevant. Thus the position of the maximum in Fig. 2 is shifted from Δ_e to $\Delta_{e\parallel}$, which is usually much smaller, resulting in a considerable increase in the frictional force for small ion velocities in the orbiter frame. The force component associated with this phenomenon is called the magnetic cooling force.

In the transverse direction one usually measures the cooling time, whereas in the longitudinal direction the cooling force can be determined directly. This is achieved by abruptly detuning the electron velocity v_0 by a small amount δ after equilibrium is reached. The ion-beam velocity is then changed to the value $v_0 + \delta$ as a consequence of the cooling force. The time taken by the ion beam to reach the new velocity is directly proportional to the longitudinal cooling force. This is illustrated in Fig. 3b.

The effect of the longitudinal force on an uncooled beam is to reduce the momentum spread. Figure 4 shows consecutive scans (at intervals of 1 s) of the momentum distribution of a proton beam immediately after injection. The initial momentum spread of 0.3% is reduced within a few seconds by a factor of about 100. Close inspection of the narrow peak thus produced (inset to Fig. 4) yields another interesting observation²⁷: at a very low beam temperature, the ions can no longer be considered as independent. The energy of the particles in the rest frame becomes comparable to the mutual Coulomb force, and collective motions can build up easily. These collective fluctuations can be represented as two waves moving with or against the direction of the beam²⁹: one group of particles moves faster and another moves slower than the nominal velocity of the ions, resulting in the two peaks seen in the inset to Fig. 4. It is predicted that the temperature of the particle beam can become so low that the energy of the particle in the rest frame becomes lower than the Coulomb force at a distance corresponding to the average distance between ions³⁰. In this case, the ions could no longer pass each other, and would circulate around the ring in a configuration that resembles a crystal, held together by their

TABLE 1 Smallest values of ion-beam parameters obtained by electron cooling

Beam radius	10 μm
Beam divergence	0.1 mrad
Beam emittance	0.1 π mm mrad
Transverse temperature	1,000 K
Momentum spread (r.m.s.)	1×10^{-6}
Longitudinal temperature	1 K
Shortest cooling time	25 ms
Highest cooling force	1 eV m ⁻¹
Lowest beam energy (protons)	1.5 MeV
Highest beam energy (protons)	287 MeV
Cooled ions	p, H ⁻ , $\bar{\text{p}}$, Li ⁺ , C ⁶⁺ , C ⁵⁺ , O ⁸⁺ , O ⁷⁺ , O ⁶⁺ , Si ¹⁴⁺

mutual electrical interactions. Indications of linear ordering have been reported^{31,32} whereby the protons move around the ring in a formation analogous to a string of pearls. If it indeed exists, this effect should be stronger, at a given temperature, for highly charged heavy-ion beams, and more complex structures have been predicted³³⁻³⁶. The experimental realization of this effect and its unambiguous theoretical proof are challenges for the new generation of cooling experiments proposed in heavy-ion storage rings.

As can be seen from Fig. 3b, the friction force can become very strong, and its effect can be seen after only a single passage. The Novosibirsk group has shown that, after a single pass, the cooling force at its maximum is about 2.5 times stronger for negative ions (H⁻) than for protons³⁷. This is because the trajectories of magnetically bound electrons are different when they scatter from a positive rather than a negative ion. This result has yet to be confirmed by comparing the cooling of protons and antiprotons.

For a long time it was widely held that electron cooling could not be applied to highly charged ions because the cooling electrons would be captured during the cooling process and the 'down-charged' ions that result would leave the storage ring. Electron capture was indeed observed in the initial proton-cooling experiments, yielding a relativistic beam of atomic hydrogen. This beam was used to measure the properties of the proton beam in a non-destructive way. The formation rate of hydrogen was of the order of a few thousand atoms per second, as predicted by theory^{5,24,38}; this loss rate is almost negligible. The recombination time τ_{rec} (the time taken to convert a fraction 1/e of the beam to down-charged ions) decreases as $1/Z_i^2$ whereas the cooling rate decreases as A/Z_i^2 (see equation (3)),

FIG. 2 A typical electron cooling device^{25,26}. Electrons are extracted from the hot cathode and accelerated in the gun to the desired velocity. They drift in a longitudinal magnetic field to the cooling section. The merging with and separation from the ion beam are achieved with toroid magnets. After the cooling process, electrons are decelerated and captured in a collector. Vacuum pumps and diagnostic elements are distributed over the whole system.

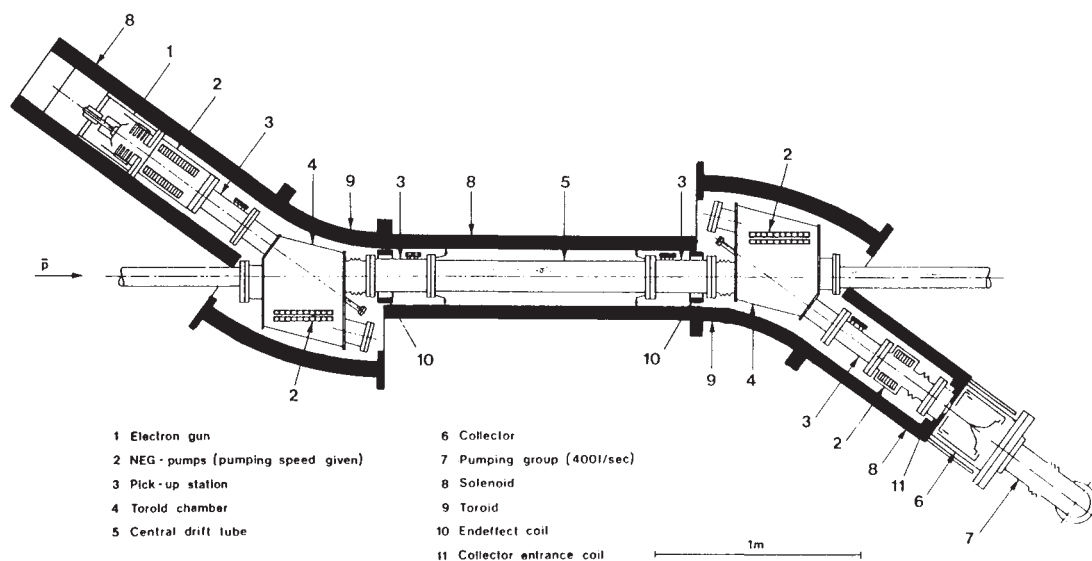
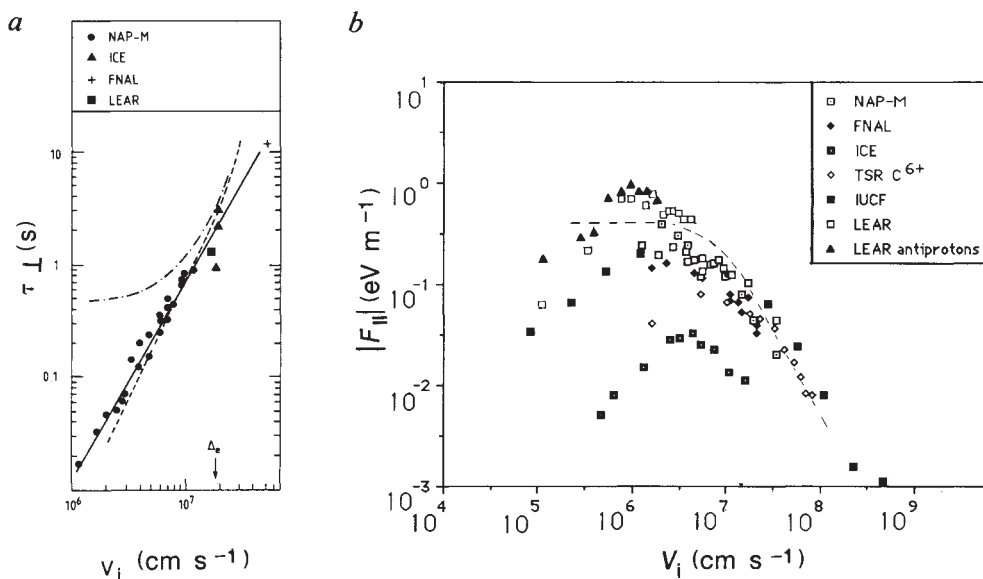


FIG. 3 Summary of electron cooling measurements for several different instruments. *a*, Transverse cooling times as a function of the ion velocity in the moving reference frame. Dashed-dotted line indicates the prediction of equation (3) (normalized to $n_e = 10^8 \text{ cm}^{-3}$ and $C/I = 50$), the dashed and solid lines are fits to the data with a model including the magnetic force. *b*, Longitudinal frictional force as a function of longitudinal ion velocity in the moving frame. The dashed line indicates the contribution from the non-magnetic force.



so the ratio τ_{rec}/τ is proportional to the nuclear mass A . For bare uranium ions the cooling time is expected to be at least a hundred times shorter than the recombination time, so that recombination is not significant over many times the cooling period τ .

Electron cooling of a 'stationary beam' of antiprotons has been carried out recently³⁹; antiprotons in a Penning ion trap were cooled from a few keV to below 1 eV. Cooling times of the order of 10 s were determined.

Some of the state-of-the-art values from electron cooling experiments^{5-9,27,28,40} are summarized in Table 1; for more details, see ref. 21.

Cooling ring projects

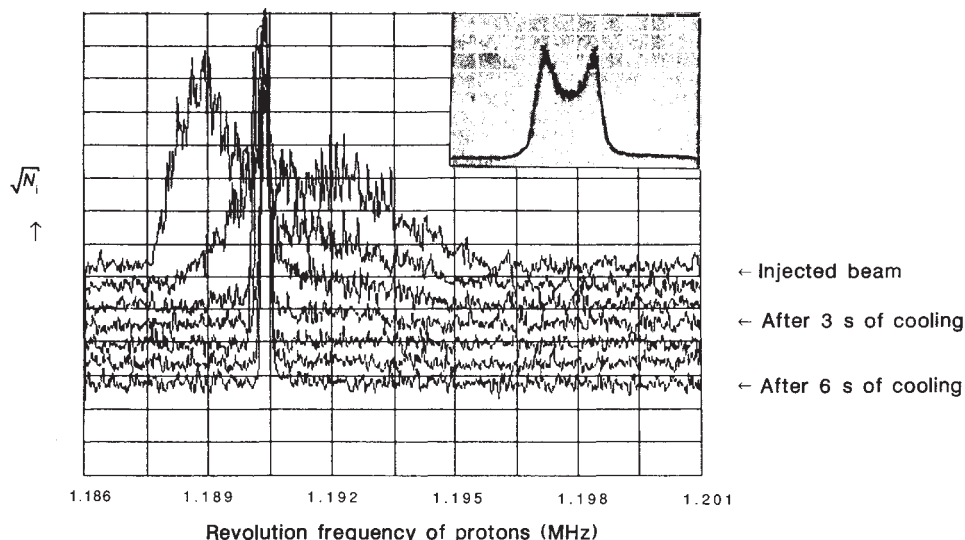
The exciting initial results of electron cooling have triggered the development of a number of storage rings for light and heavy ions equipped with coolers⁴¹⁻⁴⁹. These will provide a variety of ions and will operate over a wide range of energy (Fig. 5). The main applications of electron cooling in these instruments are summarized in Box 2. I shall discuss here some important examples.

When antiprotons are scattered from a polarized proton target, they may annihilate more frequently in one spin state than in another. This fact can be used to filter out one spin state and

thus to produce a polarized antiproton beam⁵⁰. A pure polarized proton target may be obtained for this purpose by directing a beam of atomic hydrogen through the vacuum chamber of the storage ring. But because such an atomic-beam target is very thin, the antiproton beam has to pass through it many times to become polarized. Therefore the antiprotons that do not annihilate on traversing the atomic beam are recirculated. Electron cooling is needed to counteract multiple (electromagnetic) scattering in the antiproton beam, which can otherwise lead to beam blow-up over many cycles. Gradual depletion of one spin component through annihilation would result ultimately in a polarized antiproton beam circulating in the storage ring. Such a beam has not yet been produced, but would be extremely valuable for the study of the spin dependence of the antinucleon-nucleon force and of the annihilation.

Another application is in pion production. When a projectile collides with a target nucleus, a pion can be produced if the energy available in the centre-of-mass system exceeds a threshold corresponding to the pion rest mass. The production of a charged pion converts a proton into a neutron (π^+ production) or a neutron into a proton (π^- production). The nucleons are internally bound differently, resulting in a production threshold that depends on the final nuclear states^{51,52}. A negative pion can become attached to a nuclear fragment created in the same

FIG. 4 Longitudinal cooling (cooling of momentum spread) of a stored proton beam in LEAR, showing scans of the momentum distribution immediately after injection of protons at 1-s intervals²⁷. The inset shows the final proton signal at the lowest temperature.



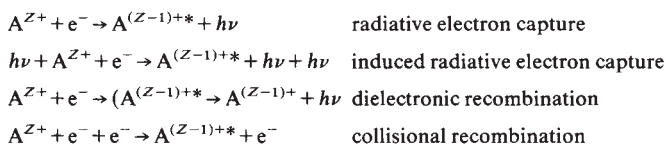
Box 2 Application of electron cooling in storage rings

Small beam sizes:	Experiments with high spatial resolution and high luminosity for colliding or merged beams.
Low momentum spread:	High-resolution nuclear and particle spectroscopy.
Long beam lifetime:	Internal-target and storage-ring operation, and beam polarization.
Fast cooling times:	Accumulation of rare ions and unstable isotopes.
High phase-space density:	Efficient deceleration of antiprotons and rare ions for high-resolution atomic spectroscopy.
Low beam temperatures:	Searches for beam ordering and crystallization.

reaction, forming a pionic atom; this introduces additional energy structure (corresponding to the atomic levels) at the threshold energy. To resolve this structure, high energy resolution is needed. The target must be thin to avoid energy smearing of the projectile and of the reaction products. These objectives can be attained using electron cooling and an internal gas target. The thickness of the target is chosen such that multiple scattering blow-up of the ion beam is counteracted by electron cooling. High-resolution energy scans through the threshold region can be performed by incremental changes in the energy of the electron beam, which drags the ion beam accordingly. The energy of the ions can be deduced with high precision from the electron acceleration voltage.

An intriguing possibility is the production of cooled beams of bare heavy ions and unstable nuclei. The short cooling times that can be obtained for highly charged ions allow sufficient quantities to be accumulated for nuclear and atomic spectroscopy. The prospect of producing dense beams of fully stripped heavy ions is appealing also for accelerator physics, because many effects of interest increase with powers of Z_i , so that highly charged ions may provide access to regions that are not easily reached with singly charged ions. Ion-beam crystallization is one such example.

Electrons used for cooling also constitute a very clean target for the study of electron-ion interactions at very low relative energy. The most important reactions of this kind are



where A denotes an ion of mass A and the asterisk refers to an excited atomic state. Processes such as these are important in plasmas and stellar atmospheres. The configuration used in electron cooling (for example, well defined electron energy, no windows, no binding effects) allows their study in heavy-ion storage rings under very clean conditions. The electron beam merged with the cooled ion beam in the interaction region represents a stable two-component plasma that is fully ionized despite its low temperature; this situation is otherwise difficult to produce and maintain. The possibilities for atomic physics are wide-ranging⁵³. One important area of investigation is the high-resolution spectroscopy of photons emitted following electron capture. A precise measurement of the electron binding energies in highly charged systems tests several predictions of quantum electrodynamics for large Z_i , a realm that is otherwise scarcely accessible in simple (hydrogen-like) systems. The large single-electron binding energy also permits relativistic effects to be investigated with great sensitivity⁵³. The emitted photons could be used as a tunable quasi-monochromatic source of X-rays⁵⁴ and the various electron capture processes might be used to extract the ion beam⁵⁵.

Antiproton and positron beams could be merged in the same way as electron and proton beams. With such an arrangement it would be possible to produce antihydrogen atoms by recombination⁵⁶. High-resolution atomic spectroscopy of antihydrogen would provide tests of the CPT theorem (the invariance of a system to simultaneous inversion of charge (C), parity (P) and time (T)) by comparing the hyperfine splitting of the antihydrogen ground state with that of hydrogen. The latter has been measured using a hydrogen maser to a precision of 6×10^{-13} . Such precision is not yet possible for antihydrogen, but even a precision of 10^{-6} would yield valuable information about the magnetic moment and the form factor of the antiproton.

There are many other appealing prospects for the study of atomic antimatter⁵⁶. High-resolution experiments with keV antiprotons should be feasible, permitting studies of their scattering from nucleons and their annihilation, as well as their capture into bound states. At present there exist virtually no experimental data for such processes at energies below ~ 10 MeV. Electron cooling would also permit large numbers of antiprotons to be brought simultaneously to rest and to be captured in a trap.

Discussion

Electron cooling is well tested and is generally well understood. It may now be used to explore new physics, particularly in experiments with non-relativistic or semi-relativistic ions and in the study of extreme states of plasmas.

Experiments using light- and heavy-ion beams profit immediately from electron cooling through phase-space compression, higher reaction rates, and better energy and spatial resolution. Subsecond cooling times, obtainable simultaneously for transverse and longitudinal degrees of freedom, and the attainment of very small beam sizes and momentum spreads, make possible the spectroscopy of nuclear and atomic states with high precision. The combination of electron cooling and use of internal targets will provide a powerful tool for such experiments. The prospects are promising for the deceleration of bare heavy ions to low energy, allowing them to be caught in a trap for the execution of high-resolution mass measurements or atomic spectroscopy. The same is true for radioactive nuclei and the study of their decay.

This is all to say that electron cooling could help to advance the frontiers of precision measurements in atomic, nuclear and low-energy particle physics. Electron cooling may also benefit high-energy physics. The rapid electron cooling of heavy ions after the last stripping stage could be used for the accumulation of many injector pulses in the acceleration cycle of a high-energy

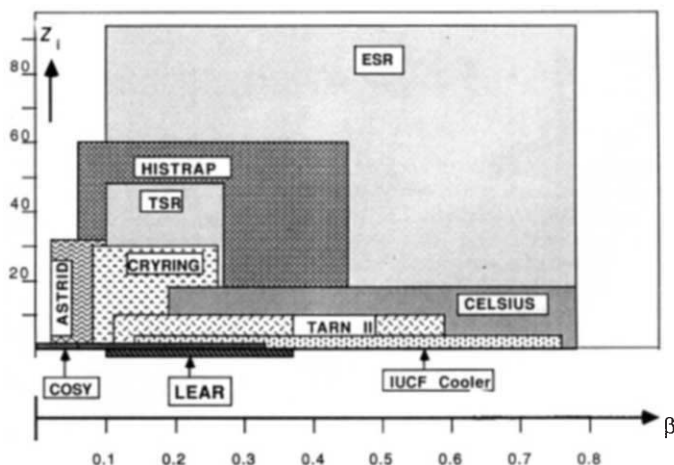


FIG. 5 Range of operation of electron cooling in storage rings; LEAR^{25,26}, IUCF^{40,41}, TSR⁴², CELSIUS⁴³, ASTRID⁴⁴, ESR⁴⁵, TARN II⁴⁶, CRYRING⁴⁷, COSY⁴⁸, HISTRAP⁴⁹. Z_i is the charge on the cooled ions, and β is the electron velocity relative to the speed of light, v_e/c .

heavy-ion accelerator. Accumulation between the cycles of a small ring used to feed a large, high-energy machine could greatly increase the beam luminosity for high-energy experiments. □

Helmut Poth is at the Gesellschaft für Schwerionenforschung, Postfach 110552, 61 Darmstadt, FRG.

1. Budker, G. I. *Proc. Int. Symp. on Electron and Positron Storage Rings* (eds Zyngier, H. & Cremieux-Alcan, E.) II-1 (Presses Univ. de France, Paris, 1967).
2. Van der Meer, S. *CERN/ISR-PO/72-31* (CERN, Geneva, 1972).
3. Möhl, D., Petrucci, G., Thorndahl, L. & Van der Meer, S. *Phys. Rep.* **58**, 73-119 (1980).
4. Van der Meer, S. *Rev. mod. Phys.* **57**, 689-697 (1985).
5. Budker, G. I. *et al. Partical Accelerators* **7**, 197-211 (1976).
6. Bell, M., Chaney, J., Herr, H., Krienen, F. & Petrucci, G. *Nucl. Instrum. Meth.* **190**, 237-255 (1981).
7. Ellison, T. *et al. IEEE Trans. Nucl. Sci.* **NS-30**, 2636-2638 (1983).
8. Poth, H. *et al. Nucl. Instrum. Meth.* **A287**, 328-332 (1989).
9. Steck, M. *et al. Nucl. Instrum. Meth.* **A287**, 324-327 (1989).
10. Blasche, K. & Böhne, D. *Proc. 14th Int. Conf. High Energy Accelerators* Chicago (American Institute of Physics, New York, 1989).
11. Derbenev, Ya. & Skrinsky, A. N. *Particle Accelerators* **8**, 1-20 (1977).
12. Derbenev, Ya. & Skrinsky, A. N. *Particle Accelerators* **8**, 235-243 (1978).
13. Ogino, T. & Ruggiero, A. G. *Particle Accelerators* **10**, 197-205 (1980).
14. Bell, J. S. & Bell, M. *Particle Accelerators* **11**, 233-238 (1981).
15. Sørensen, A. & Bonderup, E. *Nucl. Instrum. Meth.* **215**, 27-54 (1983).
16. Budker, G. I. & Skrinsky, A. N. *Soviet Phys. Usp.* **21**, 277-296 (1978).
17. Derbenev, Ya. & Skrinsky, A. N. *Phys. Rev. Vol. 3* (ed. Khalatnikov, I. M.) 165 (Harwood Academic, Chur, 1981).
18. Cole, F. T. & Mills, F. E. *A. Rev. Nucl. Particle Sci.* **31**, 295-325 (1981).
19. Skrinsky, A. N. & Parkhomchuk, V. V. *Soviet J. Particle Nucl.* **12**, 223-247 (1981).
20. Poth, H. *Proc. CERN Accelerator School*, CERN 87-03, 534-569 (CERN, Geneva, 1987).
21. Poth, H. *Phys. Rep.* (in the press).
22. Poth, H. (ed.) *Proc. Workshop on Electron Cooling and Related Applications* (ECool84) (KfK 3846, Karlsruhe, 1985).
23. Spitzer, L. *Mon. Not. R. astr. Soc.* **100**, 396-405 (1940).
24. Spitzer, L. *Physics of Fully Ionized Gases* (Interscience Publishers, New York, 1956).
25. Habfast, C. *et al. Phys. Scripta* **T22**, 277-281 (1988).
26. Poth, H. *et al. Proc. 4th Workshop on Physics at LEAR with Cooled Low Energy Antiprotons* (eds Amsler, C. *et al.*) 121 (Harwood Academic, Chur, 1988).

27. Poth, H. *et al. Z. Phys.* **A332**, 171-188 (1989).
28. Wolf, A. *et al. Proc. 1st European Accelerator Physics Conf.* (ed. Tazzari, S.) 205 (World Scientific, Singapore, 1989).
29. Parkhomchuk, V. V. & Pestrikov, D. V. *Soviet Phys. Tech. Phys.* **25**, 818-822 (1980).
30. Dikansky, N. S. & Pestrikov, D. V. *Proc. workshop on Electron Cooling and Related Applications* (ECool 84) 275 (KfK 2846, Karlsruhe, 1985).
31. Dementev, E. N., Dikansky, N. S., Medvedko, A. S., Parkhomchuk, V. V. & Pestrikov, D. V. *Soviet Phys. Tech. Phys.* **25**, 1001-1003 (1980).
32. Parkhomchuk, V. V. *Proc. Workshop on Electron Cooling and Related Applications* (ECool 84) 71 (KfK 3846, Karlsruhe, 1985).
33. Schiffer, J. P. & Kienle, P. *Z. Phys.* **A321**, 181 (1985).
34. Rahman, A. & Schiffer, J. P. *Phys. Rev. Lett.* **57**, 1133-1136 (1986).
35. Habs, D. *Frontiers of Particle Beams*, (eds Month, M. & Turner, S.) 310 (Springer, Berlin, 1986).
36. Hasse, R. W., Hofmann, I. & Liesen, D. (eds) *Proc. Workshop on Crystalline Ion Beams* (GSI-Report GSI-89-10) (GSI, Darmstadt, 1989).
37. Dikansky, M. S. *et al. Proc. 1st European Accelerator Physics Conf.* (ed. Tazzari, S.) 529 (World Scientific, Singapore, 1989).
38. Bell, M. & Bell, J. S. *Particle Accelerators* **12**, 49-52 (1982).
39. Gabrielse, G. *et al. Phys. Rev. Lett.* **63**, 1360-1363 (1989).
40. Friesel, D. L. *et al. Nucl. Instrum. Meth.* **B40/41**, 927-933 (1989).
41. Pollock, R. E. *IEEE Trans. Nucl. Sci.* **NS-30**, 2056-2060 (1983).
42. Jaeschke, E. *et al. Proc. 1st European Accelerator Physics Conf.* (ed. Tazzari, S.) 365 (World Scientific, Singapore, 1989).
43. Ekström, C. *et al. Phys. Scripta* **T22**, 256-268 (1988).
44. Stensgaard, R. *Phys. Scripta* **T22**, 315-317 (1988).
45. Franke, B. *Nucl. Instrum. Meth.* **B24/25**, 18-25 (1987).
46. Noda, A. *et al. IEEE Trans. Nucl. Sci.* **NS-32**, 2684-2688 (1985).
47. Herliander, C. J., Bárány, A., Rensfelt, K.-G. & Starker, J. *Phys. Scripta* **T22**, 282-285 (1988).
48. Pfister, U. *Proc. 1st European Accelerator Physics Conf.* (ed. Tazzari, S.) 398 (World Scientific, Singapore, 1989).
49. Olsen, D. K. *et al. Nucl. Instrum. Meth.* **B24/25**, 26-37 (1987).
50. Steffens, E. *et al. Proc. 3rd LEAR Workshop on Physics with Cooled Low-Energy Antiprotons in the ACOL Era* (eds Gastaldi, U., Klapisch, R., Richard, J.-M. & Tran Thanh Van, J.) 245 (Editions Frontieres, Gif-sur-Yvette, 1984).
51. Ericson, T. E. O. *Proc. Workshop on Nuclear Physics with Stored, Cooled Beams* (eds Schwandt, P. & Meyer, H. O.) 172 (American Institute of Physics, New York, 1985).
52. Kilian, K. *Proc. Workshop on Nuclear Physics with Stored, Cooled Beams* (eds Schwandt, P. & Meyer, H. O.) 319 (American Institute of Physics, New York, 1985).
53. Datz, S., Andersen, L. H., Briand, J.-P. & Liesen, D. *Phys. Scripta* **T22**, 224-227 (1988).
54. Poth, H. & Wolf, A. *Phys. Lett.* **A94**, 135-138 (1983).
55. Poth, H. *Z. Phys.* **A326**, 483-486 (1987).
56. Poth, H. & Wolf, A. (eds) *Proc. Symp. on the Production and Investigation of Atomic Antimatter* (ANTIMATTER '87) (Scientific Publishing Company, Basel, 1988).

Calibration of the ^{14}C timescale over the past 30,000 years using mass spectrometric U-Th ages from Barbados corals

Edouard Bard*, Bruno Hamelin*, Richard G. Fairbanks & Alan Zindler

Lamont-Doherty Geological Observatory, Columbia University, Palisades, New York 10964, USA

Uranium-thorium ages obtained by mass spectrometry from corals raised off the island of Barbados confirm the high precision of this technique over at least the past 30,000 years. Comparison of the U-Th ages with ^{14}C ages obtained on the Holocene samples shows that the U-Th ages are accurate, because they accord with the dendrochronological

calibration. Before 9,000 yr BP the ^{14}C ages are systematically younger than the U-Th ages, with a maximum difference of $\sim 3,500$ yr at $\sim 20,000$ yr BP. The U-Th technique thus provides a way of calibrating the radiocarbon timescale beyond the range of dendrochronological calibration.

RECENTLY a suite of coral cores has been raised off the island of Barbados, and more than 30 *Acropora palmata* samples were dated by conventional ^{14}C techniques¹. The analyses showed that the last deglaciation started at $\sim 15,000$ ^{14}C yr BP (before present), and that it occurred in two steps—observations in full agreement with ^{18}O records in deep-sea cores dated by accelerator mass spectrometry (AMS) of ^{14}C (refs 2-4). The

Barbados offshore collection of corals is also ideally suited for addressing two important problems: the timing of the last glacial cycle and the calibration of the ^{14}C timescale back to 30 kyr BP.

It is now well documented that radiocarbon dating is not a totally accurate chronometer because the atmospheric $^{14}\text{C}/\text{C}$ reference level has changed with time. The first evidence of a discrepancy between true age and ^{14}C age, $\Delta^{14}\text{C}$, was reported by De Vries⁵, and the most accurate evidence of changes of the $^{14}\text{C}/\text{C}$ ratio in the atmosphere (expressed as ‰ above a modern standard, ref. 6) comes from the dating of tree rings⁷. Over the past 9,000 yr, the $\Delta^{14}\text{C}$ curve is characterized by a long

* Present addresses: Centre des Faibles Radioactivités, CNRS-CEA, 91198 Gif-sur-Yvette, France (E.B.) and Geosciences de l'Environnement, Université Aix-Marseille III, 13397 Marseille, France (B.H.).

decreasing trend with a maximum difference of $\sim 100\%$ at about 7,000 yr BP. This trend has been correlated with changes of the magnetic dipole which acts as a shield to the production of ^{14}C by the cosmic rays. Whether or not this slow change is the result of variations in the ^{14}C -reservoir exchange rate rather than changes of the cosmogenic production is still being debated⁸⁻¹¹. Superimposed on the long trend are oscillations¹²⁻¹⁴, which are attributed to short-term variations of the magnetic properties of the solar wind^{13,14}. Before 9,000 yr BP the ^{14}C timescale is not accurately calibrated because of the lack of suitable fossil trees for dendrochronological dating. Several authors have attempted to extend the calibration using other means of establishing absolute ages: varved sediment^{15,16}, U-Th¹⁷⁻¹⁹ and ice cores²⁰. The results obtained are not very precise and are often in disagreement with one another (up to 1,000%). Nevertheless it seems that before 9,000 yr BP the ^{14}C ages are consistently younger than the true ages, suggesting higher $^{14}\text{C}/\text{C}$ ratios than present (a positive $\Delta^{14}\text{C}$).

Recently ^{234}U - ^{230}Th dating of corals by thermal ionization mass spectrometry (TIMS) has been demonstrated²¹⁻²⁵ to be significantly more precise and accurate than the α -counting method. It has also been predicted that U-Th dating would be even more precise than the ^{14}C in the time range of the last deglaciation (6-20 kyr). We have implemented this technique at the Lamont-Doherty Geological Observatory and measured Th and ^{14}C ages in the same coral samples to enable comparison of the two geochronometers. Here we focus on the U-Th/ ^{14}C comparison for the past 30,000 years. The implications of the data for the timing of the last glacial cycle as well as other U-Th ages obtained on older corals (back to 130,000 yr) will be described elsewhere (E.B., B.H. and R.G.F., manuscript in preparation).

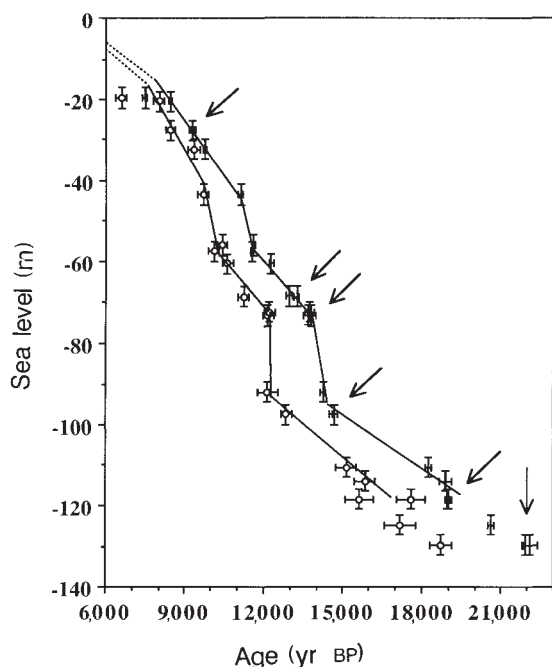


FIG. 1 Sea level during the last deglaciation. Crosses correspond to U-Th ages and the open symbols to ^{14}C ages (calculated with the 5,730-yr half-life). Highest and two lowest samples are composed of *Porites asteroides* and thus only represent lower bound for the sea level at those times. All other samples are *Acropora palmata*. The age errors are quoted at 2σ . The samples marked by arrows correspond to duplicate analyses of the same coral sample by U-Th mass spectrometry. All duplicates are in agreement within the limits of error. A linear uplift correction has been applied by assuming that the last interglacial highstand was at 7 m above the present sea level. This correction is of the order of the inherent error due to *A. palmata* habitat range (± 2.5 m).

Methods

All isotope measurements were performed on a VG Micromass 30, single-collector mass spectrometer, equipped with an analog Daly detector for the analysis of small-abundance isotopes. The techniques used for chemical separation and mass spectrometry of U and Th have been adapted from ref. 22. They will be described in detail elsewhere (E.B., B.H., R.G.F. and A.Z., manuscript in preparation).

For thorium we used a ^{229}Th spike for measurement of ^{230}Th by isotope dilution. The absolute concentration of this spike was calibrated against a dilution of a SPEX ThO_2 standard. For the corals measured so far, the $^{232}\text{Th}/^{230}\text{Th}$ ratios are generally lower than 30 and therefore no correction for detrital ^{230}Th is required because the typical crustal $^{232}\text{Th}/^{230}\text{Th}$ values are of the order of 10^5 (ref. 23). For uranium we used a ^{236}U spike with 99% enrichment for the measurement of ^{238}U and ^{234}U abundances. The absolute concentration of this spike was calibrated against a dilution of NBL CRM111 ^{233}U standard. The accuracy of our correction procedures for isotope fractionation and mass 238 contribution was checked using the NBS U-010 solution: for the $^{234}\text{U}/^{238}\text{U}$ atomic ratio the weighted mean of six measurements is $(5.43 \pm 0.02) \times 10^{-5}$ in agreement with the NBS-certified isotope ratio of $(5.466 \pm 0.06) \times 10^{-5}$.

Results

Table 1 gives the list of U-Th results; as an initial check of the accuracy and precision of our dating technique we dated samples collected from the uplifted reefs of Barbados and Haiti, which correspond to the last interglacial period (samples AFM3, C4 and C1 cf Table 1). These measurements agree with previous α -counting results²⁶ and confirm that Th ages measured in samples from the last interglacial highstand are restricted to the time between 120 and 130 kyr BP (ref. 23).

The precision of the U-Th dating is typically better than 100 yr (at 2σ) for ages between 6 and 20 kyr, significantly better than that achieved for most ^{14}C dating. In addition, replicate analyses of the same hand samples demonstrate that the method is reproducible within the 2σ errors (nine duplicate couples; see Table 1). Initial $^{234}\text{U}/^{238}\text{U}$ ratios calculated using the Th ages fall in the range of modern sea water (1.14-1.15; ref. 21) for all deglaciation samples. This shows that their U-Th systematics have not been strongly altered after deposition.

Figure 1 focuses on the coral record of the last deglaciation dated by U-Th and ^{14}C . The implication of this plot on the chronology and structure of the last deglaciation is described elsewhere (E.B., B.H. and R.G.F., manuscript in preparation). It is remarkable that there is not even a single stratigraphic inversion in the U-Th data. This shows that erosion and redeposition of coral fragments are minor problems in our Barbados core collection.

Samples RGF9-27-5 (dated at Geochron) and RGF9-27-4 (dated at Lamont-Doherty) correspond to the same depth in core 9 (5 cm from each other) but the difference in ^{14}C age (1,900 yr) is very large. The ^{14}C date at 17 kyr BP is considered to be anomalous as it produces a stratigraphic inversion in the core-9 sequence (see Table 1 and ref. 1).

The measurements of ^{14}C in corals cannot be used directly for calibration against absolute chronometers because the surface layer of the ocean is not in equilibrium with the atmosphere; instead, a steady-state balance is maintained between the input of ^{14}C from the atmosphere and its removal by advection and radiodecay in the water column. The ^{14}C ages obtained on corals must therefore be corrected for the age difference between the atmosphere and the surface ocean. This parameter depends mainly on the upwelling and mixing with old sub-surface water masses. For the low-latitude range the reservoir age is generally ~ 400 yr (50%) and is not likely to have changed by more than 100-200 yr during the late Pleistocene²⁷. We have therefore adjusted the raw ^{14}C ages by 400 yr in Fig. 1 and the following discussion.

TABLE 1 U/Th and ^{14}C results from Barbados corals

Sample	Species	Depth (m)	^{238}U (p.p.m.)	$^{230}\text{Th}/^{234}\text{U}$	$^{234}\text{U}/^{238}\text{U}$	$(^{234}\text{U}/^{238}\text{U})_{\text{init}}$	Th age (yr BP)	^{14}C age (yr BP)	$^{14}\text{C}_{5,730}$ (yr BP)	^{14}C lab
RGF7-4-2	Pa	-18	2.65 ± 0.02	0.0667 ± 0.0007	1.1414 ± 0.0065	1.1444 ± 0.0067	$7,460 \pm 80$	$6,400 \pm 200$	6,586	L-DGO
RGF7-5-5	Ap	-19	3.24 ± 0.02	0.0752 ± 0.0004	1.1468 ± 0.0052	1.1504 ± 0.0053	$8,450 \pm 50$	$7,780 \pm 220$	8,006	Geochron
RGF7-12-2#1	Ap	-25	3.16 ± 0.02	0.0823 ± 0.0008	1.1359 ± 0.0079	1.1395 ± 0.0081	$9,280 \pm 100$	$8,200 \pm 200$	8,439	L-DGO
RGF7-12-2#2			3.31 ± 0.02	0.0820 ± 0.0007	1.1455 ± 0.0088	1.1494 ± 0.0091	$9,250 \pm 80$			
RGF7-16-2	Ap	-30	3.19 ± 0.01	0.0861 ± 0.0004	1.1456 ± 0.0040	1.1497 ± 0.0041	$9,730 \pm 50$	$9,050 \pm 250$	9,313	Geochron
RGF7-27-4	Ap	-41	3.45 ± 0.02	0.0975 ± 0.0006	1.1411 ± 0.0061	1.1456 ± 0.0063	$11,090 \pm 70$	$9,400 \pm 200$	9,673	L-DGO
RGF12-5-2	Ap	-53	3.17 ± 0.02	0.1016 ± 0.0005	1.1407 ± 0.0043	1.1454 ± 0.0045	$11,590 \pm 60$	$10,100 \pm 200$	10,394	L-DGO
RGF12-6-7	Ap	-55	3.04 ± 0.02	0.1011 ± 0.0006	1.1463 ± 0.0062	1.1512 ± 0.0064	$11,530 \pm 70$	$9,800 \pm 200$	10,085	L-DGO
RGF12-9-5	Ap	-58	3.04 ± 0.01	0.1072 ± 0.0007	1.1462 ± 0.0056	1.1514 ± 0.0058	$12,260 \pm 90$	$10,300 \pm 200$	10,600	L-DGO
RGF12/16/05#1	Ap	-65	2.96 ± 0.01	0.1151 ± 0.0009	1.1399 ± 0.0048	1.1453 ± 0.0050	$13,220 \pm 110$	$10,900 \pm 200$	11,217	L-DGO
RGF12/16/05#2			3.02 ± 0.01	0.1130 ± 0.0010	1.1402 ± 0.0071	1.1455 ± 0.0074	$12,970 \pm 120$			
RGF12-21-6	Ap	-69	3.03 ± 0.02	0.1190 ± 0.0014	1.1446 ± 0.0064	1.1503 ± 0.0067	$13,700 \pm 170$	$11,850 \pm 200$	12,195	L-DGO
RGF12/21/10#1	Ap	-70	3.33 ± 0.02	0.1198 ± 0.0011	1.1335 ± 0.0057	1.1388 ± 0.0060	$13,800 \pm 140$	$11,800 \pm 200$	12,143	L-DGO
RGF12/21/10#2			3.21 ± 0.02	0.1186 ± 0.0011	1.1395 ± 0.0107	1.1450 ± 0.0112	$13,660 \pm 140$			
RGF9-8-2	Ap	-89	3.06 ± 0.02	0.1233 ± 0.0008	1.1372 ± 0.0069	1.1429 ± 0.0072	$14,230 \pm 100$	$11,800 \pm 400$	12,143	L-DGO
RGF9-13-3#1	Ap	-94	3.22 ± 0.02	0.1267 ± 0.0013	1.1309 ± 0.0058	1.1365 ± 0.0061	$14,660 \pm 160$	$12,500 \pm 200$	12,864	L-DGO
RGF9-13-3#2			3.22 ± 0.02	0.1271 ± 0.0008	1.1314 ± 0.0061	1.1370 ± 0.0064	$14,700 \pm 100$			
RGF9-21-11	Ap	-106	3.52 ± 0.03	0.1552 ± 0.0011	1.1358 ± 0.0039	1.1430 ± 0.0042	$18,240 \pm 140$	$14,700 \pm 400$	15,128	L-DGO
RGF9/24/04	Ap	-109	3.36 ± 0.02	0.1603 ± 0.0019	1.1346 ± 0.0079	1.1420 ± 0.0084	$18,890 \pm 250$	$15,400 \pm 400$	15,848	L-DGO
RGF9-27-5#1	Ap	-114	3.43 ± 0.01	0.1610 ± 0.0007	1.1363 ± 0.0050	1.1439 ± 0.0053	$18,980 \pm 90$	$17,085 \pm 520$	17,582	Geochron
RGF9-27-5#2			3.45 ± 0.02	0.1614 ± 0.0008	1.1391 ± 0.0055	1.1468 ± 0.0058	$19,030 \pm 100$			
RGF9-27-4	Ap	-114						$15,200 \pm 400$	15,642	L-DGO
RGF9-32-4	Pa	-120	3.15 ± 0.02	0.1735 ± 0.0009	1.1329 ± 0.0048	1.1409 ± 0.0051	$20,610 \pm 120$	$16,700 \pm 600$	17,186	L-DGO
RGF9-34-8#1	Pa	-124	3.22 ± 0.01	0.1836 ± 0.0011	1.1335 ± 0.0041	1.1421 ± 0.0044	$21,930 \pm 150$	$18,200 \pm 400$	18,730	L-DGO
RGF9-34-8#2			3.16 ± 0.02	0.1851 ± 0.0019	1.1349 ± 0.0094	1.1437 ± 0.0101	$22,130 \pm 260$			
RGF12-30-2#1	Ap	-78	3.22 ± 0.02	0.2458 ± 0.0016	1.1255 ± 0.0067	1.1368 ± 0.0074	$30,470 \pm 240$	$27,120 \pm 1,520$	27,909	Geochron
RGF12-30-2#2			3.46 ± 0.02	0.2428 ± 0.0014	1.1314 ± 0.0062	1.1431 ± 0.0068	$30,040 \pm 210$			
AFM3#1	Ap	55	3.04 ± 0.02	0.6944 ± 0.0046	1.1077 ± 0.0072	1.1536 ± 0.0110	$125,100 \pm 1,700$			
AFM3#2			3.12 ± 0.01	0.6941 ± 0.0026	1.1065 ± 0.0037	1.1518 ± 0.0057	$125,100 \pm 1,000$			
Haiti C4#1	Ap	52	2.95 ± 0.01	0.6897 ± 0.0043	1.1094 ± 0.0072	1.1553 ± 0.0109	$123,600 \pm 1,600$			
C4#2			2.95 ± 0.01	0.6937 ± 0.0035	1.1071 ± 0.0051	1.1526 ± 0.0078	$124,900 \pm 1,300$			
Haiti C1	Ap	52	3.10 ± 0.01	0.6843 ± 0.0030	1.1070 ± 0.0050	1.1512 ± 0.0075	$121,900 \pm 1,100$			

Depth or altitude of recovery are relative to the present sea level. Ap and Pa refer respectively to *Acropora palmata* and *Porites asteroideus* samples. The measured atomic ratios have been converted to activity ratios by using the same U-Th half-lives used in ref. 22 (all errors are given by 2σ). Labels #1 and #2 correspond to duplicate U/Th analyses. The initial $^{234}\text{U}/^{238}\text{U}$ ratios are obtained by using the Th ages to correct the measured isotope ratios. L-DGO is the Lamont-Doherty Geological Observatory. Note that there is no significant change of uranium concentration between glacial and Holocene samples. The ^{14}C ages correspond to conventional ages (calculated with the 5,568-yr half-life) and ages calculated with the true 5,730-yr half-life. The ^{14}C ages have been corrected for fractionation and a reservoir age of 400 yr (ref. 27).

The most striking result is that the Th ages are consistently older than the ^{14}C ages, with a maximum difference of ~ 3.5 kyr at ~ 20 Th kyr BP (Fig. 2). The difference between the two chronometers is equivalent to a $\Delta^{14}\text{C}$ of ~ 400 – 500% above the present-day reference level for ages between 18 and 22 kyr BP.

Comparison with previous ^{14}C calibration

The three youngest ^{14}C -dated coral samples (RGF7-4-2, RGF7-5-5 and RGF7-12-2, Table 1) from the Barbados collection are in the range of the precise tree-ring calibration. Our U-Th ages for these samples are in agreement with the dendrochronological ages (7,440–7,140, 8,970–8,380 and 9,420–8,970 yr BP are the calibrated ^{14}C dates to be compared with the Th ages of samples RGF7-4-2, RGF7-5-5 and RGF7-12-2, respectively). For the comparison we used the mixed-layer calibration curve of Stuiver *et al.*²⁸ and for the oldest sample, RGF7-12-2, we used unpublished data (M. Stuiver, personal communication). The good agreement between the two independent calibrations is a confirmation of the accuracy of the U-Th chronometer in corals, at least for the Holocene (see refs 22, 24 for the accuracy on very young corals). Before the Holocene, we cannot yet demonstrate that our U-Th ages are true 'calendar' ages, but we will argue below that this is a reasonable assumption.

Based on a comparison between varve counts and ^{14}C dating in Lake of the Clouds¹⁵, the calibration curve was extended to 11 kyr BP. There is also good agreement between the varve estimates and the U-Th/ ^{14}C comparison (Fig. 2).

From 9 to 13 kyr BP, attempts have been made to use the Swedish varve chronology to extend the ^{14}C calibration. These sediments were never directly dated by ^{14}C , however, and the Swedish chronology was tied to the present only recently²⁹. Even in its latest version the $\Delta^{14}\text{C}$ curve based on the Swedish varves correlation is still not in agreement with the Lake of the Clouds record for the interval between 8 and 11 kyr BP and also disagrees with our Barbados results. Stuiver³⁰ concluded that a few hundreds of varves are missing in the Swedish chronology.

At $\sim 10,000$ ^{14}C yr BP (Preboreal/Younger Dryas boundary) the U-Th calibration ($\sim 11,500$ Th yr BP) is in relatively good agreement with the dendrochronological estimates ($\geq 11,300$

dendro yr BP, ref. 31). These two independent estimates are in disagreement with the correlation between climate spectra dated by annual counting of seasonal variations in the ^{18}O content, dust and acidity curves from the Dye III and Camp Century ice cores ($\sim 10,700$ yr BP, ref. 20).

A detailed analysis of the U-Th/ ^{14}C calibration indicates that the true duration of the Younger Dryas climate event (10,000–11,000 ^{14}C yr BP) may have been as long as 1,700 years, in agreement with the qualitative estimates by Oeschger *et al.*³². Moreover, it is not inconsistent with the plateau in ^{14}C ages at $\sim 9,500$ ^{14}C yr BP seen in the dendrochronological data of ref. 31.

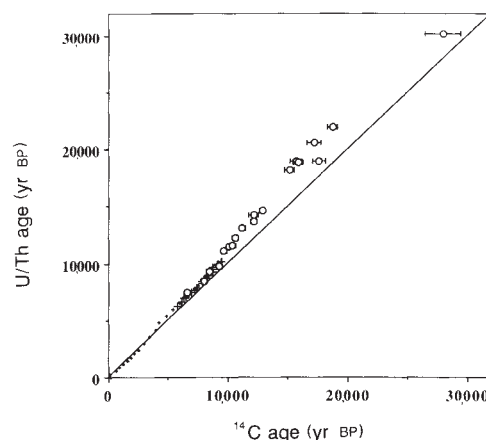


FIG. 2 Open symbols represent the U-Th ages plotted against ^{14}C ages (calculated with the 5,730-yr half-life). The small crosses correspond to the dendrochronological calibration⁷ and the large crosses to the Lake of the Clouds varved sediments¹⁵. For clarity the duplicate analyses by U-Th mass spectrometry have been averaged. The age errors are quoted at 2σ (in the case of U-Th ages they are always smaller than the size of the symbol). For sample RGF9-27-5 we have used the ^{14}C analyses of sample RGF 9-27-5 (Geochron) and RGF9-27-4 (Lamont-Doherty). This study suggests that the ^{14}C age of ~ 17 kyr is anomalous (stratigraphic inversion, see Table 1). For the three younger samples, the good agreement between the tree-ring calibration and our U-Th/ ^{14}C estimates indicate that the Th ages are accurate.

Other attempts of radiocarbon calibration have been made by comparing ^{14}C ages with U-Th ages obtained in samples from dried lakes or speleothems¹⁷⁻¹⁹. For these works the U-Th ages were obtained by α -counting with a much lower precision than by mass spectrometry (the 2σ errors are typically 10-20% of the ages) and a considerable uncertainty is associated with a ^{14}C correction for the initial age of the bicarbonate. Nevertheless, the Barbados U-Th data agree with the study on a south African speleothem¹⁹ and with the Searles Lake record^{17,18} if we use the ^{14}C reservoir age for this lake suggested by Broecker and Kaufman³³.

For the 12,000-25,000 ^{14}C yr BP period other authors have found large discrepancies between ^{14}C and K/Ar dating⁵² and between ^{14}C and thermoluminescence dating⁵³, in agreement with our Barbados results.

Possible geochemical artefacts

One explanation for the increasing lag between ^{14}C and Th ages could be that the samples were contaminated by recent carbon, either *in situ* and/or during the ^{14}C determinations. To correct the ^{14}C ages back to the Th ages, a contamination of ~4% modern carbon is required. We consider that this is highly unlikely because: (1) replicates conducted so far by AMS do not show any significant difference whether or not the coral sample is strongly leached before processing (conducted on samples RGF9-21-11, RGF12-30-2, RGF1-17-4; E.B., M. Arnold, B. H. and R.G.F., manuscript in preparation); (2) the ^{14}C ages obtained by Fairbanks¹ were measured by β -counting in two different laboratories (Lamont-Doherty, without leaching the samples; Geochron lab, with leaching of the samples); and (3) 2% of contamination is the maximum acceptable to keep

the three recent ages in agreement with the tree-ring ages. A definitive test of the accuracy of the ^{14}C ages used in this comparison will be to replicate all the measurements by AMS.

The good quality of the corals for U-Th was tested by X-ray diffraction which showed that no secondary calcite was detectable in the corals. This test is important, but not sufficient, to rule out any diagenetic alteration or uranium exchange³⁴. The strongest argument for the quality of the samples comes from the U-Th measurements themselves: all the initial $^{234}\text{U}/^{238}\text{U}$ ratios fall in the range of the modern sea water (activity ratio between 1.14 and 1.15, ref. 21) in contrast with older air-exposed samples (Table 1, see also ref. 22 and B.H., E.B. and R.G.F., manuscript in preparation).

A gradual loss of uranium could also explain the discrepancy between the two geochronometers. This dissolution can be modelled with first-order kinetics³⁴. A rate of uranium loss equivalent to $k = 7\lambda_{234}$ (where λ_{234} is the decay constant of ^{234}U) would be required to reconcile the ^{14}C ages with the U-Th ages. This explanation seems unlikely, however, because a strong decrease in the uranium concentration with time should be evident in the record (at ~30 kyr BP the decrease should be 50%). The data in Table 1 demonstrate that the uranium concentration in the corals is actually rather constant (3-3.5 p.p.m.), in agreement with other published estimates for *A. palmata*²².

Carbon cycle changes

The ^{14}C -age calculation relies on the assumption that the atmospheric $^{14}\text{C}/^{12}\text{C}$ ratio has not changed with time and therefore can be used to normalize the measured ratio in fossil carbon³⁵. Unfortunately, the atmosphere contains less than a ton of ^{14}C and most of the radiocarbon is in the ocean (>40 tons). Subtle changes in the reservoir sizes and/or rates of exchange between reservoirs could have significantly affected the ^{14}C content of the atmosphere. CO_2 measurements in old polar ice showed that the atmospheric concentration was lower during the last glacial than during the Holocene³⁶⁻³⁸. Box-model calculations show that the steady-state $\Delta^{14}\text{C}$ would be increased by 25-75%

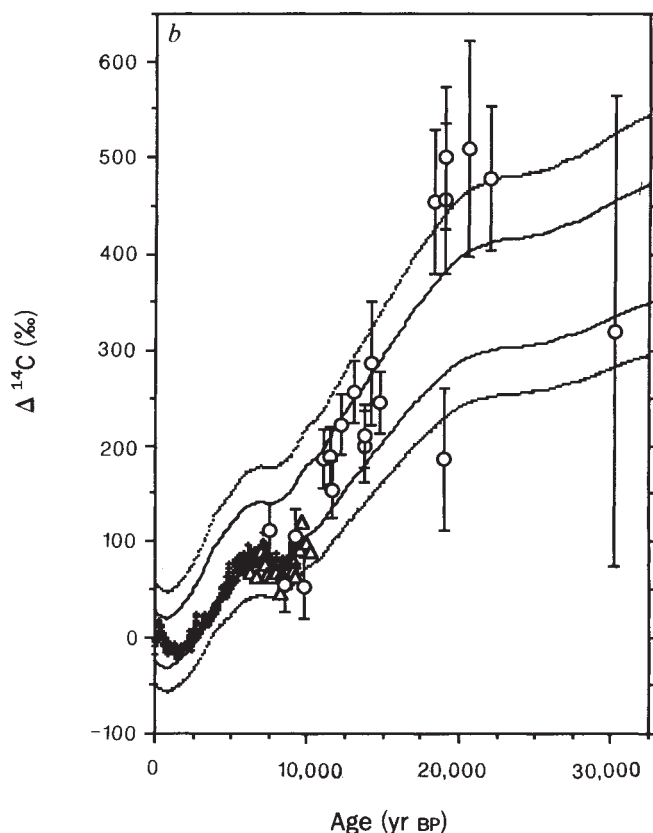
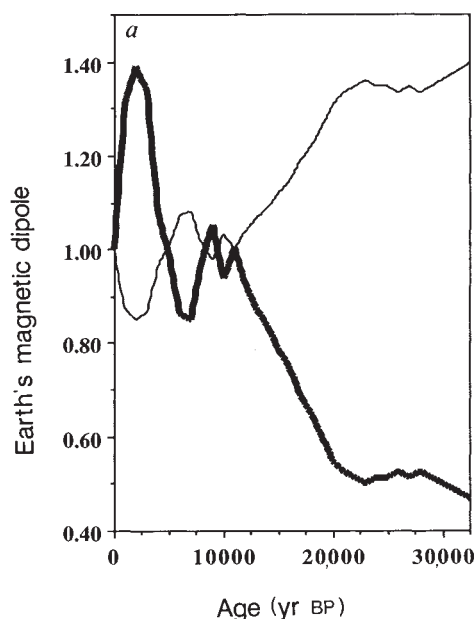


FIG. 3 a, Thick line, intensity of Earth's magnetic dipole reconstructed through time (adapted from ref. 48); thin line, cosmogenic nuclide production through time calculated using these palaeomagnetism data and theoretical calculations by Lal⁵⁰. The scale is normalized to unity for present-day values. Using b, atmospheric $\Delta^{14}\text{C}$ through time reconstructed by the U-Th/ ^{14}C age relationship (open symbols, the errors are quoted at 2σ). The small crosses correspond to the tree-ring calibration⁷ and the open triangles to the Lake of the Cloud calibration (ref. 15, no error has been represented). The two sets of solid lines correspond to the envelopes of $\Delta^{14}\text{C}$ expected as a response to changes of the Earth's magnetic field: for those calculations we have assumed constant errors of ± 5 and $\pm 10\%$ of the present-day dipole. For sample RGF 9-27-5 we have used the ^{14}C analyses of sample RGF9-27-5 (Geochron) and RGF9-27-4 (Lamont-Doherty). This study suggests that the ^{14}C age of ~17 kyr is anomalous (stratigraphic inversion, see Table 1).

depending on the glacial-to-interglacial CO_2 variation and on the version of the box model used^{39–41}.

It is possible to increase further the atmospheric $\Delta^{14}\text{C}$ by diminishing the rate of penetration of $^{14}\text{CO}_2$ in the deep sea. But, to increase the atmospheric $^{14}\text{C}/\text{C}$ by 400% it would be necessary to change dramatically the relative ^{14}C age of the deep waters. By contrast, the mixed layer of the ocean is rapidly exchanging with the atmosphere. A model calculation similar to that of Oeschger *et al.*⁴² shows that the difference in ^{14}C between the surface and the deep sea would have to be increased by ~ 3.5 kyr to account for a 400% higher atmospheric $^{14}\text{C}/\text{C}$ (ref. 27). This age difference is not supported by the palaeoceanographic data on benthic-planktonic pairs of foraminifera dated by ^{14}C AMS. These studies have shown that the ventilation age difference was higher in the Pacific by $\sim 500 \pm 100$ yr (refs 43, 44) and in the Atlantic by 200 ± 200 yr (ref. 44). Assuming that the highest values measured in the Pacific are valid for the global ocean would lead to a maximal atmospheric $^{14}\text{C}/\text{C}$ ratio increase of $\sim 70\%$.

In summary, changes in the carbon cycle can account for only ~ 100 – 150% increase in the $\Delta^{14}\text{C}$ of the atmosphere as compared to the 400–500% inferred from the Barbados data.

Changes in the ^{14}C production

The mean $^{14}\text{C}/\text{C}$ ratio of the exchangeable carbon on the Earth is maintained by a steady state between production by cosmic rays in the upper atmosphere and radiodecay. The $\Delta^{14}\text{C}$ of the atmosphere can therefore undergo variations in response to changes in the flux of galactic protons, solar-wind magnetic properties or the dipole moment of the Earth.

Little is known about possible variations of the cosmic rays on the timescale of interest. As discussed below, the main $\Delta^{14}\text{C}$ variations during the Holocene can be explained by changes of the magnetic field and/or changes in the ocean circulation. Therefore it is not necessary to call for modulation of the galactic protons, and this flux can be assumed to be constant with time until data is produced to the contrary.

Changes in the solar-wind modulation are also poorly known but it is important to mention that Raisbeck *et al.*⁴⁵ have described brief periods of enhanced cosmogenic production during the past 80 kyr, which they attribute to solar effects. Their work, based on AMS measurements of ^{10}Be in the Vostok ice core, shows that the strongest event occurred at ~ 35 and 60 kyr BP. In general, the ^{10}Be record obtained in ice cores shows consistently higher concentrations during glacial stages than during warmer periods (a factor of two for the last deglaciation), but the generally accepted interpretation is that precipitation was strongly reduced during the cold periods⁴⁶. Consequently, aside from the solar ^{10}Be spikes, it is extremely difficult to separate the climate signal and the cosmogenic-production signal embedded in the measurements of ^{10}Be in ice cores.

What is well documented, however, is the fact that the magnetic moment of the Earth has varied significantly over the past several thousand years, as first shown by Thellier and Thellier⁴⁷. The palaeointensity record for the past 40 kyr is characterized by two important features: a quasi-sinusoidal variation during the Holocene and a decrease of the moment by a factor of two before 20 kyr BP (refs 48, 49 and Fig. 3a).

We converted this dipole history into a ^{14}C production time series by using the calculations of Lal (ref. 50; Fig. 3a). The atmospheric ^{14}C production curve cannot be compared directly with the $\Delta^{14}\text{C}$ record, however, because the ^{14}C atoms mix continuously in other carbon reservoirs where they ultimately decay with a mean half-life of 8,270 yr. The carbon cycle therefore has a long 'memory', which can easily be simulated by using a simple box model^{39,51}. The carbon cycle acts as a low-pass filter: short-term production variations (such as the 11-yr solar cycle) are smoothed out and a phase shift of a few kyr is introduced for the transition between the Pleistocene and the Holocene^{39,51}. The results of the calculation are shown in Fig.

3b. We assumed no error on the ^{14}C production algorithm⁵⁰ and constant errors of 5 and 10% of the present-day magnetic field on the palaeomagnetic record. There is clear agreement between the $\Delta^{14}\text{C}$ data calculated from the U-Th/ ^{14}C comparison and our calculation based on the palaeomagnetic record. The $\Delta^{14}\text{C}$ values around 18–22 kyr BP are, however, higher on average than the palaeomagnetism reconstruction, which may indicate that $\sim 100\%$ of the signal is the result of reservoir changes associated with the last glacial maximum.

It must be mentioned that some of the palaeomagnetism data compiled by McElhinny and Senanayake⁴⁸ were dated by ^{14}C . Using our calibration (Fig. 2) to correct the timescale would only stretch the predicted $\Delta^{14}\text{C}$ curve by a few kyr. The U-Th/ ^{14}C would still be well within the errors of the geomagnetic reconstruction of the $\Delta^{14}\text{C}$ curve. Therefore the most reasonable explanation for the large difference between the ^{14}C and U-Th ages involves magnetic modulation of the atmospheric ^{14}C production rate.

Discussion

We have shown here that by using U-Th mass spectrometry one can achieve an age precision and reproducibility comparable to or even better than that obtained by the radiocarbon geochronological method. U-Th ages measured in corals drilled off the shore of Barbados indicate that the ^{14}C timescale is significantly compressed during the late glacial interval. These results can be explained mainly as the consequences of changes in the cosmogenic nuclide production linked to the gradual decrease of Earth's magnetic field⁴⁸. Changes in the carbon cycle^{36–39,43,44} can account for ~ 10 – 20% of the age discrepancy observed between the two geochronometers. It should be borne in mind that the accuracy of the U-Th ages is proven only back to 10,000 years; nevertheless, all geophysical information available (dendrochronology, varve dating and palaeomagnetism) suggests that our U-Th results are reasonable and that they may be used as a first-order tool to calibrate the radiocarbon timescale beyond the range of the high-resolution dendrochronological comparison. □

Received 16 January; accepted 20 April 1990.

1. Fairbanks, R. G. *Nature* **342**, 637–642 (1989).
2. Duplessy, J. C. *et al.* *Nature* **320**, 350–352 (1986).
3. Bard, E. *et al.* *Nature* **328**, 791–794 (1987).
4. Broecker, W. S. *et al.* *Nature* **333**, 156–158 (1988).
5. De Vries, H. L. *Proc. Koninkl. Ned. Akad. Wetenschap.* **B61**, 94–102 (1958).
6. Stuiver, M. & Polach, H. A. *Radiocarbon* **19**, 355–363 (1977).
7. Stuiver, M. & Kra, R. *Radiocarbon* **28**, 805–1030 (1986).
8. Lal, D. *Geophys. Monogr. Ser.* **32**, 221–233 (1985).
9. Beer, J. *et al.* *Nature* **331**, 675–679 (1988).
10. Andree, M. *et al.* *Clim. Dyn.* **1**, 53–62 (1986).
11. Peng, T. H. *Radiocarbon* (in the press).
12. Suess, H. in *Proc. Conf. Problems Related to Interplanetary Matter* NAS-NSF Publ. 845, Nucl. Sci. Ser. Vol. **33**, 90–95 (NAS-NSF, La Jolla, 1961).
13. Stuiver, M. *J. geophys. Res.* **66**, 273–276 (1961).
14. Stuiver, M. & Quay, P. D. *Science* **207**, 11–19 (1980).
15. Stuiver, M. in *Late Cenozoic Glacial Ages* (ed. Turekian, K. K.) 57–70 (Yale Univ. Press, New Haven, 1970).
16. Tauber, H. in *Nobel Symp. 12th, Radiocarbon Variations and Absolute Chronology* (ed. Olsson, I. U.) (Wiley, New York, 1970).
17. Peng, T. H., Goddard, J. G. & Broecker, W. S. *Quat. Res.* **9**, 319–329 (1978).
18. Stuiver, M. *Nature* **273**, 271–274 (1978).
19. Vogel, J. C. *Radiocarbon* **25**, 213–218 (1983).
20. Hammer, C. U., Clausen, H. B. & Tauber, H. *Radiocarbon* **28**, 284–291 (1986).
21. Chen, J. H., Edwards, R. L. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **80**, 241–251 (1986).
22. Edwards, R. L. thesis, California Inst. Technol. (1988).
23. Edwards, R. L., Chen, J. H. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **81**, 175–192 (1987).
24. Edwards, R. L., Taylor, F. W. & Wasserburg, G. J. *Earth planet. Sci. Lett.* **90**, 371–381 (1988).
25. Li, W. X. *et al.* *Nature* **337**, 534–536 (1989).
26. Dodge, R. E., Fairbanks, R. G., Benninger, L. K. & Maurrasse, F. *Science* **219**, 1423–1425 (1983).
27. Bard, E. *Paleoceanography* **3**, 635–645 (1988).
28. Stuiver, M., Pearson, G. W. & Braziunas, T. *Radiocarbon* **28**, 980–1021 (1986).
29. Cato, I. *Boreas* **14**, 117–122 (1985).
30. Stuiver, M., Kromer, B., Becker, B. & Ferguson, C. W. *Radiocarbon* **28**, 969–979 (1986).
31. Becker, B. & Kromer, B. *Radiocarbon* **28**, 961–967 (1986).
32. Oeschger, H. *et al.* *Radiocarbon* **22**, 299–310 (1980).
33. Broecker, W. S. & Kaufman, A. *Geol. Soc. Am. Bull.* **76**, 537–566 (1965).
34. Ku, T. L., Ivanovitch, M. & Luo, S. *Quat. Res.* (in the press).
35. Libby, W. F. *Radiocarbon Dating* (Univ. Chicago Press, 1952).
36. Delmas, R. J., Ascencio, J. M. & Legrand, M. *Nature* **284**, 155–157 (1980).
37. Berner, W., Oeschger, H. & Stauffer, B. *Radiocarbon* **22**, 227–235 (1980).
38. Barnola, J. M., Raynaud, D., Korotkevitch, Y. S. & Lorius, C. *Nature* **329**, 408–414 (1987).

39. Siegenthaler, U., Heimann, M. & Oeschger, H. *Radiocarbon* **22**, 177–191 (1980).
40. Keir, R. S. *Earth planet. Sci. Lett.* **64**, 445–456 (1983).
41. Lal, D. & Revelle, R. *Nature* **308**, 344–346 (1984).
42. Oeschger, H., Siegenthaler, U., Schotterer, U. & Gugelmann, A. *Tellus* **27**, 168–192 (1975).
43. Shackleton, N. J. *et al. Nature* **335**, 708–711 (1988).
44. Broecker, W. S. *et al. Paleoceanography* **3**, 659–669 (1988).
45. Raisbeck, G. M. *et al. Nature* **326**, 273–277 (1987).
46. You, F., Raisbeck, G. M., Bourles, D., Loric, C. & Barkov, N. I. *Nature* **316**, 616–617 (1985).
47. Thellier, E. & Thellier, O. *C. r. hebdom. Seanc. Acad. Sci. Paris* **212**, 281–283 (1941).
48. McElhinny, M. W. & Senanayake, W. E. *J. Geomagn. Geoelec.*, Kyoto **34**, 39–51 (1982).
49. Tauxe, L. & Valet, J. P. *Phys. Earth planet. Inter.* **56**, 59–68 (1989).
50. Lal, D. in *Proc. 45th Conf. Solar-Terrestrial Relationships and the Earth Environment in the Last Millennium*, 216–233 (Soc. Italiana di Fisica, Bologna, 1988).
51. Houtermans, J. C., Suess, H. E. & Oeschger, H. *J. geophys. Res.* **78**, 1897–1908 (1973).
52. Gillot, P. Y. & Cornette, Y. *Chem. Geol.* **59**, 202–222 (1986).
53. Valladas, H. & Valladas, G. *Coll. Int. CNRS, Habitats du Paléolithique Supérieur* (Roanne-Villerey, in the press).

ACKNOWLEDGEMENTS. We thank G. Mathieu for the ^{14}C determinations at Lamont-Doherty and W. S. Broecker for discussions and support. We acknowledge B. Anderson, M. Arnold, J. C. Duplessy, L. Froget, C. Laj, Y. Lao, W. X. Li, J. Lin, J. Rubenstone and D. Schneider for discussions and/or technical support. This research was supported by the NSF and NASA.

The identification and suppression of inherited neurodegeneration in *Caenorhabditis elegans*

Martin Chalfie & Eve Wolinsky*

Department of Biological Sciences, 1012 Fairchild, Columbia University, New York, New York 10027, USA

The dominant mutation *deg-1(u38)* results in a toxic gene product that leads to the late-onset degeneration of a small number of neurons in the nematode *Caenorhabditis elegans*. Both intragenic and extragenic mutations as well as changes in wild-type gene dosage can delay or block the time of onset of the neuronal deaths. The *deg-1* gene has been cloned and a partial complementary DNA reveals that the gene encodes a novel protein that may act as a membrane receptor. Because the late-onset loss of specific sets of neurons, often as a result of dominant mutations, is characteristic of several human neurodegenerative diseases, the analysis of the *deg-1* gene and its suppressors may provide a means of understanding the mechanisms underlying some of these human diseases.

THE selective degeneration of specific classes of nerve cells is characteristic of many inherited human disorders¹ such as Huntington's disease^{2,3}, familial amyotrophic lateral sclerosis⁴, familial Alzheimer's disease⁵ and several cerebellar ataxias⁶. Many of these diseases are caused by dominant alleles that result in late-onset degeneration of the affected nerve cells, but their molecular bases are not known. Insights into how genetic lesions can lead to neuronal degeneration could be obtained by studying animal models with similar defects.

We have characterized a dominant mutation (*u38*) of the *C. elegans* gene *deg-1* (degeneration), that results in the degeneration of a small set of neurons. One striking feature of some of these deaths is that they have a late onset, occurring long after the neurons form synapses and become functional. The *deg-1(u38)* mutation results in an abnormal, toxic form of a gene product that is normally non-essential for neuron viability and function. The affected cells swell to many cell diameters and lyse. Because the *deg-1* gene seems to encode a membrane component, cell lysis may result from compromised membrane integrity. The time of the neuronal degeneration depends upon the level of *deg-1(u38)* mutant gene activity, suggesting that the level of accumulation of toxic product, rather than the time of its initial expression, causes the deaths. Mutations in the *mec-6* gene suppress the *deg-1(u38)* neurodegeneration, as well as

similar neuronal deaths produced by abnormal toxic forms of the *mec-4* gene product. An understanding of how the mutant *deg-1* and *mec-4* gene products result in neurodegeneration, and of how normal *mec-6* (mechanosensory) activity is required for this degeneration might help in the development of methods for preventing or curing neurodegenerative disorders in humans.

The *deg-1(u38)* phenotype

The *deg-1(u38)* mutant was identified in a screen for touch-insensitive mutants following ethyl methanesulphonate (EMS) mutagenesis at 25 °C (ref. 7). The gene is located on the X chromosome (see legend to Fig. 1). The *deg-1(u38)* animals differ from other touch-insensitive mutants^{7,8} in that although both types of mutants are insensitive to the gentle touch of a hair, only the *deg-1* mutants are insensitive to the more severe prod of a thin wire. They are also insensitive only at the tail. This more severe phenotype (designated Tab for Touch abnormal) suggests a defect in the two PVC interneurons, which receive synaptic inputs from posterior touch receptor neurons⁹. The PVC cells degenerate in these mutants. A similar behavioural abnormality results from laser ablation of the PVC cells in wild-type animals (W. W. Walthall and M.C., unpublished data).

Although the PVC cells arise during embryogenesis¹⁰, *deg-1* mutants are touch-sensitive at hatching and become Tab later, during the second and third larval stages (L2 and L3) at 25 °C (Fig. 1). The stage of onset of touch insensitivity is delayed further when animals are grown at lower temperatures, so mutants grown at 15 °C become Tab as gravid adults.

Virtually all *deg-1* animals, when viewed by Nomarski microscopy¹¹, have one or two degenerating cells near the normal position of the PVC cells, 24 h after hatching at 25 °C (Fig. 2). The degeneration is first seen as a small vacuole that surrounds the nucleus. This vacuole, which contains particles displaying Brownian motion, enlarges by several cell diameters over the next few hours, during which the nucleus disintegrates.

This degeneration differs from programmed cell death, a common feature in *C. elegans* development, in which affected cells become refractile and condensed as they die¹¹. Here we use the term degeneration for the type of death seen in *deg-1(u38)* mutants to distinguish it from the morphologically distinct programmed cell death. The *deg-1(u38)* degeneration phenotype (termed Deg) has been seen in animals with dominant mutations of the gene *mec-4* (*mec-4(d)*). Dominant, but not recessive, *mec-4* mutations result in the degeneration of the touch receptor neurons^{7,8}, probably by the production of a toxic product that is expressed within the touch receptor neurons¹². As in *deg-1(u38)* animals, some of the affected cells in *mec-4(d)* mutants display a late-onset degeneration^{8,13}.

* Present address: Department of Biochemistry, New York University Medical School, 550 First Avenue, New York, New York 10016, USA.

Other classes of neurons, unrelated by position, lineage, or function to each other or to the PVC cells, also degenerate at various times during development in *deg-1(u38)* mutants (Fig. 2). Although all of the animals become Tab and most have one or two degenerating PVC cells at 24 h at 25 °C, the pattern of these other deaths varies considerably (Fig. 2 legend). This variability may reflect incomplete penetrance of the Deg phenotype for each cell or variations in the times of onset or duration of the degenerations.

The *deg-1* mutants also contain empty vacuoles (vacuoles not associated with nuclei) at the tip of the nose, along the ventral cord, and in the tail (Fig. 2), especially in newly hatched animals. They may be located in the hypodermis (epidermis), but whether they result from *deg-1* expression in this tissue or are a response to neuronal damage is unknown.

Genetic characterization

Alleles like *u38* seem to be extremely rare, as no similar *deg-1* mutation has yet been isolated. The original *u38* mutant was isolated in a screen of progeny representing 70,000 copies of the gene⁷. Such rarity (loss-of-function mutations in *C. elegans* arise at a frequency of $\sim 5 \times 10^{-4}$ (refs 14, 15)) and the dominant Deg and Tab phenotypes suggest that *u38* is not a *deg-1* null mutation. Indeed, loss of *deg-1* function gives rise to no detectable abnormal phenotype (the animals appear wild type) and the *u38* mutation seems to encode an abnormal product that interferes with normal cell activity.

The *u38* animals can readily be reverted to a wild-type phenotype. As 24 of 25 non-Tab (touch-sensitive) revertants obtained by EMS mutagenesis have not segregated Tab animals, these reverting mutations are presumed to be intragenic. (The remaining revertant strain produces some Tab progeny but with an unusual segregation pattern and has not been studied further.) Three additional intragenic revertants arose spontaneously from

a *mut-2; deg-1* double mutant (the *mut-2* mutation often activates transposons in *C. elegans*¹⁶). Most of the suppressor mutations, including those obtained in the *mut-2* background, seem to revert almost completely the Tab and Deg phenotypes, and all but one are recessive to the *u38* mutation. The frequency at which the intragenic suppressor mutations arose with EMS mutagenesis (4×10^{-4}) is similar to that found for loss-of-function mutations in other *C. elegans* genes^{14,15}. The molecular analysis below also supports the hypothesis that disruption of the *deg-1* gene leads to a wild-type null phenotype.

The high frequency of intragenic *deg-1* revertants that do not produce a detectable phenotype suggests that the wild-type *deg-1* product is not essential. The gene could be nonessential because of a marginal effect of *deg-1(+)* in the wild type or the replacement of its function by the products of other genes, either of the same gene family^{15,17} or of a redundant genetic pathway (for example, ref. 18). Whereas the loss of a constituent of a gene family or one pathway might not result in observable defects, a mutation leading to the production of an abnormal, interfering gene product could. The molecular analysis below suggests that a *deg-1* gene family exists, but the possibility of a redundant pathway cannot be excluded.

Two of the intragenic suppressor mutations, *u352* and *u354*, only partially suppress the effects of *u38*. They delay the onset of the Tab phenotype, by about 24 and 48 h at 25 °C for *u354* and *u352*, respectively (Fig. 3). The PVC deaths are also seen later, the cells dying in young and old adults, respectively. The various deaths caused by the *u38* mutation may be unequally affected by these suppressor mutations; in particular the *u352* mutation seems to have its most pronounced effect on the PVC death (Fig. 3).

Gene dosage experiments suggest that the *u38* mutation results in an abnormal gene product that interferes with cell function, rather than in the overproduction of a normal cell product. Additional copies of the wild-type gene delay, although to a smaller degree than temperature or intragenic mutations, the onset of the *u38* Tab phenotype (Fig. 4). In general, the delay in the onset of the Tab phenotype increases in proportion to the ratio of wild-type to mutant allele.

Independence of *deg-1(u38)* degenerations

The *deg-1* degenerations differ from programmed cell deaths not only morphologically, but also genetically. Mutations that prevent all programmed cell deaths in *C. elegans* hermaphrodites (*ced-3(n171)* and *ced-4(n1162)*; ref. 19) did not prevent the *deg-1* degenerations. In fact, *ced-3; deg-1(u38)* and *ced-4; deg-1(u38)* double mutants had additional degenerations in the head at hatching (total deaths with *ced-3*: 4.0 ± 0.4 ; with *ced-4*: 4.6 ± 0.6 ; mean $\pm$ s.e.m., $n = 25$ for each; *deg-1(u38)* mutants alone have 2.2 ± 0.2 deaths, $n = 20$). These data suggest that some of the extra cells that normally survive in *ced-3* and *ced-4* animals live and then degenerate in the corresponding *deg-1* double mutant. This supports the hypothesis^{19,20} that cells prevented from undergoing programmed cell death can differentiate, in this case at least enough to express the Deg phenotype.

One explanation for the late-onset degeneration of the PVC cells is that larval cell interactions trigger the deaths, perhaps by initiating aberrant differentiation or by toxically stimulating the cells, as in glutamate neurotoxicity²¹. Many neurons arise post-embryonically in *C. elegans*, but they do not seem to be required for the PVC deaths, as the double mutant *lin-6(e1466); deg-1(u38)*, in which none of these cells are made^{22,23}, is Tab. Moreover, the elimination of the touch receptor neurons, which synapse onto the PVC cells, by the addition of the *unc-86(e1416)* mutation⁸ had no effect in the *deg-1* deaths. Thus, neither postembryonically derived targets nor a major input to the PVC cells is required for these deaths. Together with the different timing and position of the degenerations, these data suggest that the *deg-1(u38)* mutation may act in a cell-autonomous fashion, although they do not rule out the possibility that cell interactions

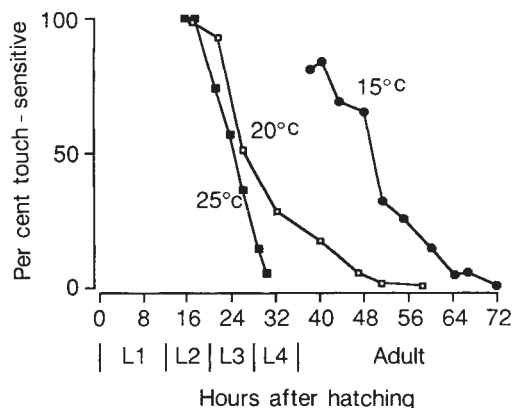


FIG. 1 Onset of the *deg-1* touch-insensitivity (Tab) phenotype at various temperatures. The abscissa records time in 25 °C-hour equivalents (*C. elegans* grows at approximately twice the 15 °C rate at 25 °C).

METHODS. Strains were maintained on *Escherichia coli* strain OP50-1 as described^{14,56}. Animals from strains grown at the indicated temperatures for at least two generations were synchronized at hatching by washing hatched animals and bacteria from the plates with M9 buffer¹⁴ and collecting the larvae that hatched from the remaining eggs in the next 1–2 h. At the indicated times animals were tested for touch sensitivity at the tail by touching with a thin hair⁸. At least 100 animals were examined at each temperature. The *deg-1(u38)* mutation maps within 0.014 map units of the *mec-7* on the X chromosome by the following criteria: no wild-type progeny were found among 3,533 total progeny from $+/+deg-1(u38)/lon-2(e678)$ [*mec-7(e1506)+*] heterozygotes nor among 5,334 and 5,960 heterozygote progeny from, respectively, $+/+deg-1(u38)$ [*dpy-6(e14)/dpy-7(e88)*] [*mec-7(e1506)+*] and *dpy-7* [$+/+deg-1(u38)$]/ $+/+[mec-7(e1506)+]$ [*dpy-6(e14)*] heterozygotes. These other mutations are described in refs 11 and 14. *deg-1* loss-of-function mutations complement the recessive *mec-7* allele *e1506* for touch insensitivity, an indication that the two loci are different genes.

could trigger a *deg-1*-dependent degeneration. Herman¹² has suggested the *mec-4(d)* mutation also acts cell-autonomously.

Suppression by *mec-6* mutations

Mutations in the *mec-6* gene suppress the degenerations caused by both the *mec-4(d)* and *deg-1(u38)* mutations. This result not only indicates an underlying genetic similarity in the deaths caused by these mutations, but also identifies a gene, *mec-6*, whose activity is required for the neurodegenerative process. The effect on the *mec-4(d)* mutations was discovered through the construction of a set of double mutants, each containing a *mec-4(d)* mutation and a mutation in one of the other genes required for touch receptor function⁷ to determine whether any of these mutations could prevent the *mec-4(d)* deaths. (Previous experiments (ref. 8, and unpublished results) showed that mutations in genes required for the generation of the touch receptors or for the specification of their differentiation prevent the

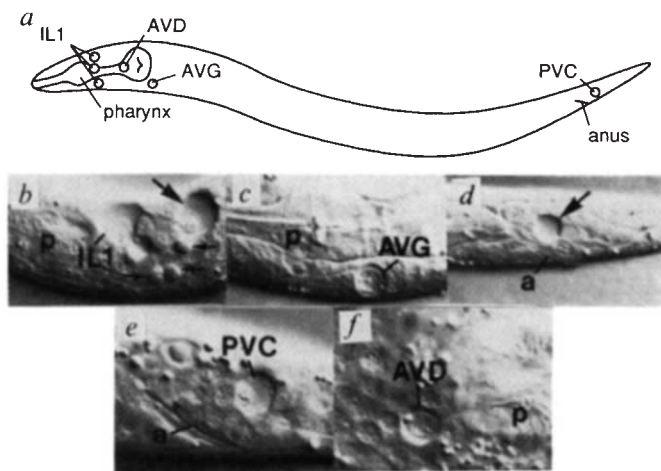


FIG. 2 Defects in *deg-1* mutants. *a*, Schematic diagram of the left side of newly hatched *C. elegans* larva indicating the positions of some of the cells affected by the *deg-1(u38)* mutation. All of these neurons, except AVG, have homologues on the other side of the animal. These cells die at different times: the IL1 sensory neurons and the AVG interneuron die at hatching, the PVC cells in the third larval stage (L3), and the AVD interneurons near the end of the last (L4) larval stage. Other degenerations are sometimes seen. Because the degenerations can be quite large, distorting the surrounding pattern of nuclei, it is difficult to identify unambiguously the affected cells. The degeneration pattern in *u38* males does not differ substantially from that of hermaphrodites. *b*, Differential interference contrast micrographs of *deg-1* degenerative deaths (IL1 and large arrow) and programmed cell deaths (small arrows) posterior to the first bulb of the pharynx (p) in a newly-hatched *ced-1; deg-1* double mutant. (The *ced-1* mutation *e1735* delays the engulfment of cells dying by programmed cell death⁵⁷.) The two types of deaths are morphologically distinct. The cells here and in subsequent photomicrographs have been tentatively identified by criteria given below. Not all of the IL1 cells die: one to three IL1 cells (wild type have six) were seen in six *deg-1* adults examined by serial section electron microscopy^{58,59}. Magnification, $\times 800$. *c*, Putative AVG death in a newly-hatched larva. 24% (12/49) newly hatched animals had this death. *d*, Empty vacuole in the tail of a newly hatched larva. *a*, Anus. *e*, PVC death in 24 h larva. *f*, Putative AVD death at the time of the last larval moult (36 h). Criteria for identifying cells: The dying cells were identified by their positions⁵⁹ and the following data. IL1: Cells with the characteristic rootlet⁵⁹ of the IL1 cells were missing in serial sections of the tip of the nose of adult animals. AVG: A single degeneration is seen in the retrovesicular ganglion just posterior to the rear bulb of the pharynx in many newly hatched animals. AVG is the only unpaired neuron in this region⁶⁰. AVD: This pair of interneurons helps mediate anterior touch sensitivity⁹, but since the anterior touch circuit also utilizes other interneurons, loss of the AVD cells in the *deg-1* mutants does not produce a Tab phenotype. However, laser ablation of the AVA interneurons in *deg-1* animals results in animals that are virtually incapable of backward movement as young adults but not as larvae (J. Thomas, personal communication), a result seen in wild-type animals when both the AVA and AVD cells are ablated⁹. PVC, See text.

appearance of the *mec-4(d)*-dependent deaths.) Mutations in the genes *mec-1*, *mec-2*, *mec-6* to *mec-10*, *mec-12*, *mec-14*, *mec-17* and *mec-18* were tested, but only *mec-6* mutations suppressed the *mec-4(d)* degenerations. Although the animals remained touch insensitive (the *mec-6* phenotype), the touch cells did not die. In 14 *mec-6-mec-4(d)* strains made using eight *mec-6* alleles and the three *mec-4(d)* mutations, all had normal appearing adult touch cells. The *mec-6* mutations suppressed only when homozygous. The *mec-6* mutation *u247* results in a temperature-sensitive phenotype for touch sensitivity⁷. Suppression of the *mec-4(d)* degenerations by *u247* is also temperature sensitive: deaths were seen in newly-hatched larvae at 15 °C, but not at 25 °C (it was not possible to delay the onset of these deaths in temperature-shift experiments). These experiments suggest that the loss of *mec-6* activity, not the production of an unusual, allele-specific product, results in suppression. As other features of the touch receptors seem unaffected in *mec-6* mutants^{7,8}, the wild-type *mec-6* product is probably required for neurodegeneration induced by the abnormal products of the *mec-4(d)* alleles.

Mutations in *mec-6* (*u41*, *u247* and *u450* were used) also suppress the effects of *deg-1(u38)*. Double mutants are insensitive to the touch of hair at both head and tail (the *mec-6* phenotype), but they are not Tab. Moreover, the *mec-6* mutations seem to suppress all of the *deg-1*-induced deaths (no vacuolated cells were seen in any newly-hatched larvae) as well as the appearance of the presumptive hypodermal vacuoles (Fig. 2*d*). As with the suppression of the *mec-4(d)* deaths, suppression only occurred in *mec-6* homozygotes. Suppression of *deg-1* by *mec-6* indicates that *mec-6* expression is not restricted to the touch system, even though the only detectable behavioural phenotype of *mec-6* mutants is touch insensitivity. The absence of a Tab phenotype in *mec-6* mutants suggests that *deg-1* in the PVC cells is not replaced by the activity of other *deg-1*-like genes, at least not those requiring *mec-6* function.

Molecular analysis of *deg-1*

The *deg-1* gene was cloned by transposon tagging, utilizing the *mut-2*-derived revertants (Fig. 5). DNA flanking the insertion site of the transposon Tc1 hybridized to C47C12 (Fig. 6), a cosmid clone adjacent to a cosmid containing the *mec-7* gene^{24,25}, a position consistent with the genetic mapping of *deg-1*

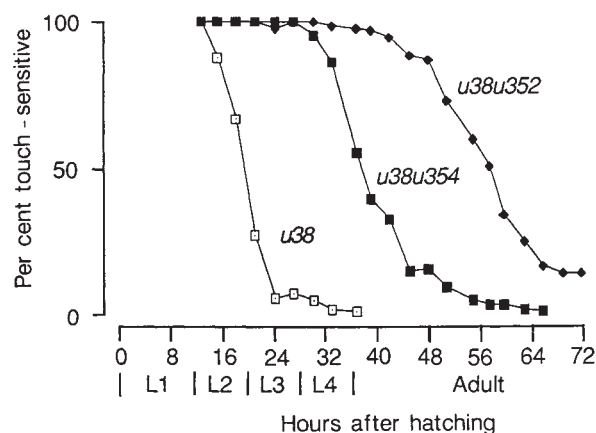


FIG. 3 Onset of the Tab phenotype in *deg-1(u38)* and two partially reverted strains [*deg-1(u38u354)* and *deg-1(u38u352)*]. Animals were grown from synchronized populations at 25 °C (see Fig. 1). At least 100 animals of each strain were examined. The Tab phenotype, as in *u38* mutants, is temperature dependent in *u38u352* and *u38u354* animals (data not shown). Not all *deg-1*-dependent deaths seem to be affected in the same way by these suppressor mutations. The number of degenerations in newly hatched animals is reduced in *u38u354* animals (0.6 ± 0.2 deaths) but not in *u38u352* animals (2.9 ± 0.3 deaths; $n=20$ for each; *u38* animals have 2.2 ± 0.2 deaths). Both suppressor-containing strains produce animals with the presumptive AVD deaths (Fig. 2) at 36 h (5/11, 3/11 and 2/10 for *u38*, *u38u352* and *u38u354*, respectively).

(Fig. 1 legend). Hybridization with C47C12 revealed insertions in one other *mut-2*-derived revertant and a possible deletion in one EMS-derived revertant (Fig. 5). Further evidence that the transposon insertions are in the *deg-1* gene came from experiments in which the transposon-containing strains were themselves reverted by reintroducing the *mut-2* mutation. This reversion with the reappearance of Deg and Tab phenotypes was accompanied by the excision of the transposon (Fig. 5).

Two partial *deg-1* cDNAs were isolated from a λ gt10 library with cDNAs from 12-hour-old larvae, provided by J. Ahlinger and J. Kimble (University of Wisconsin). As the transposon is inserted in the revertant *u38 u478* residues in a 187-bp region that contains sequences in the cDNAs surrounding a 50-bp intron (see Fig. 6 legend), it is likely that these cDNAs represent *deg-1* transcripts. The larger cDNA encodes an open reading frame of 884 base pairs (bp; Fig. 7) followed by 141 bp and a poly(A) tail. The smaller cDNA lacks 72-bp from within the open reading frame. As no untranslated sequence precedes these open reading frames, the cDNAs seem to be incomplete. The inferred polypeptide sequence of 294 amino acids from the larger cDNA, which is not strongly homologous to other DNA and protein database entries, has three notable features: (1) a hydrophobic region from residues 215–255; (2) two possible sites for *N*-linked glycosylation²⁶ located at positions 8 and 122; and (3) a cysteine-rich region, containing 13 of the 14 cysteines in the sequence, between positions 66 and 187. The 72 bp absent in the smaller cDNA encode 24 amino acids in the cysteine-rich region (including two cysteines) and could represent an alternatively-spliced variant.

The structure of the *deg-1* product is consistent with its being a membrane protein. The hydrophobic domain has an uninterrupted stretch of 17 amino acids, which is sufficient to span the lipid bilayer²⁷, and an additional 24 amino acids, only three of which are charged. This structure but not sequence, is also found in the predicted membrane spanning domain of the β subunit of the T-cell receptor^{28,29}. It is possible that the extended hydrophobic region of the *deg-1* product (including the charged residues) is submerged in the cell membrane, and may interact with analogous domains of other membrane proteins. The localization of the large number of cysteines, a feature of several receptor proteins^{30–33}, and the two *N*-linked glycosylation sites in the *N*-terminal region, is consistent with this region being an extracellular domain (Fig. 7).

The molecular analysis supports the hypothesis derived from our genetic studies that *deg-1* may be a member of a gene family. The cDNAs as well as a genomic DNA fragment to which they bind (the IC fragment; Fig. 6) hybridize to several genomic fragments (Fig. 5b).

The wild-type *deg-1* gene, as defined by the restriction fragments that differ in the *u38* revertants or that hybridize with the cDNAs, is contained within a 35 kilobase (kb) *Bam*HI fragment (Fig. 6). The corresponding fragment cloned from the dominant allele *deg-1(u38)* (Fig. 6, TU#3), when transformed into wild-type animals, produces both the Tab and Deg phenotypes (Fig. 6 legend). Most of the transformant lines are unstable and seem to have extrachromosomal arrays of the injected DNA. However, in one strain (*uIn1*) the mutant phenotypes are stably inherited, and karyotype and Southern blot analysis indicate that at least ten intact copies of the transformed DNA have attached to the X chromosome.

The *uIn1* animals resemble *u38* mutants, but display some differences. The PVC neurons degenerate at essentially the same time in both strains (data not shown), but more degenerating cells are seen in the head during the second larval stage (L2) in the transformed animals (0.5 ± 0.0 deaths in *u38* ($n = 21$) and 3.7 ± 0.2 in *uIn1* ($n = 9$)). The extra copies of *deg-1(u38)* DNA or its misexpression in the transformants may cause additional cells to die or change the onset of some cell degeneration. Unlike the situation in *u38* mutants, the Deg and Tab phenotypes of *uIn1* animals are not completely suppressed by *mec-6* mutations

u450 and *e1342*. The head degenerations in *mec-6; uIn1* strains are suppressed, but the PVC degenerations occur in many of the animals (8 of 28 late L2 *mec-6; uIn1* larvae had these deaths compared with 7 of 20 *uIn1* larvae of the same age). Most *mec-6; uIn1* adults are Tab. Hence, some *deg-1*-dependent degenerations do not absolutely require wild-type *mec-6* function, at least under these conditions of increased copies of the *u38* DNA.

Discussion

Dominant mutations of *deg-1* and *mec-4* cause the degeneration of a small number of neurons. Although different cells die in the *deg-1* and *mec-4* mutants, these mutations cause an apparently similar cell death that differs morphologically and genetically from programmed cell death. The degeneration result from the expression of alleles that seem to encode abnormal products; the wild-type genes are not needed for cell viability. The dominant mutations are suppressed by *mec-6* mutations, suggesting that similar molecular processes may underlie these neuronal degenerations. Over 420 mutations (many of them dominant) in 18 genes needed for touch receptor development and function have been identified, yet only the three *mec-4(d)* mutations cause the deaths of the touch cells⁷, and no other mutations are known that mimic the *deg-1(u38)* mutation. Perhaps only a few cellular processes can be disrupted to cause this type of selective neurodegeneration. (This is not the only type of abnormal cell death seen in *C. elegans*; mutations in *egl-1*³⁴, *lin-24*³⁵ and *lin-33*³⁵ seem to produce ectopic cell deaths.)

An intriguing feature of the *deg-1(u38)* phenotype is the late onset of cell degeneration of many of the affected cells. Because the loss of the PVC cells results in a detectable phenotype, the behavioural defect in these animals also seems to be of late onset, even though other cells have died earlier. The PVC degenerations can occur in animals of different ages (from second stage larvae to egg-laying adults) depending on the nature of the mutation, the growth temperature, and the dosage of the wild-type gene. These data suggest that the degenerative

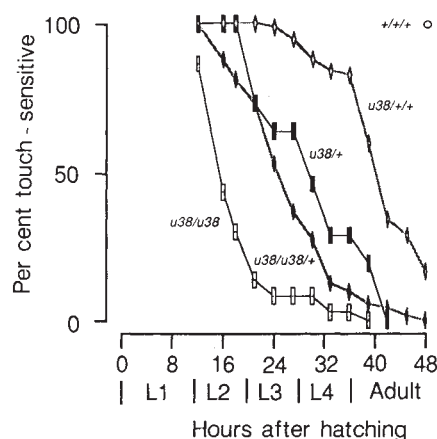


FIG. 4 Onset of the Tab phenotype in animals with different doses of the wild-type (+) and mutant (*u38*) alleles of *deg-1*. Progeny of hermaphrodites with the genotype *stDp2/+; deg-1(u38)+dpy-6(e14)/unc-18(u81)+dpy-6(e14)* were synchronized at hatching (Fig. 1 legend). At the indicated times after hatching animals that were Tab were plated individually so that their progeny could be examined so as to determine their genotypes. All animals that have three copies of the region (*u38/u38/+* (filled diamonds), *u38/+* (open diamonds) and *+/+* (open circle)) contain *stDp2*, a duplication of the *dpy-6 deg-1* region of the X chromosome that has been translocated to chromosome II and is lethal when homozygous (R. Waterston, pers. comm.). All these animals are wild type in length. At least 50 animals of each of these genotypes were examined. Animals with two copies of the region (*u38/u38* (open bars) and *u38/+* (filled bars)) are homozygous for the *dpy-6* mutation and are shorter in length (Dpy). Thirty-seven *u38/u38* and eleven *u38/+* animals were examined. The Dpy phenotype does not seem to affect the onset of the degeneration (compare Fig. 1). (The *dpy* and *unc* mutations are described in ref 16.) Wild-type animals (*+/+*; not shown), like the *+/+* animals, are touch-sensitive at all times.

phenotype is not directly coupled to a specific, developmental event, such as the appearance or maturation of a particular set of cells, a conclusion supported by the observations that the *deg-1*-dependent PVC deaths are not affected when various synaptic targets and inputs have been eliminated genetically. More likely, the accumulation of a mutant product leads to the degeneration phenotype. Certainly, a particular time of onset cannot be taken as a characteristic feature of the Tab phenotype. Such may also be true of human diseases. For example, the study of restriction fragment length polymorphisms linked to the Huntington's disease gene suggested that mutant alleles of the same genetic locus may give rise to diseases with different times of onset³⁶.

One possible explanation for the marked effect of *deg-1* and *mec-4* mutations on specific nerve cells, as well as the vacuolization phenotype, is that the genes encode membrane com-

ponents such as ion channels or receptors, that are required by subsets of neurons. An abnormal product from a dominant mutation might prove toxic if, directly or indirectly, it compromised membrane integrity. Several observations suggest that chronically open membrane channels can lead to cell lethality. The continued opening of acetylcholine channels at the neuromuscular junction, for example, leads to localized degeneration of the endplate because of the activation of calcium-dependent proteases³⁷, and sustained opening of capsaicin-sensitive channels in putative pain receptor neurons can cause cell lysis through osmotic imbalance or calcium influx³⁸. The osmotic disruption and calcium entry that result from glutamate neurotoxicity³⁹ may also result from a similar mechanism. The partial sequence of the *deg-1* product is consistent with a role in membrane function, although it is not homologous to known channel proteins. The *deg-1* product could affect membrane integrity indirectly. For example, the mutant product might bind to a channel or other membrane component and modify its activity. Of relevance may be a mammalian protein with only one apparent transmembrane domain that confers novel channel properties on *Xenopus* oocytes^{40,41}. It is not known whether this

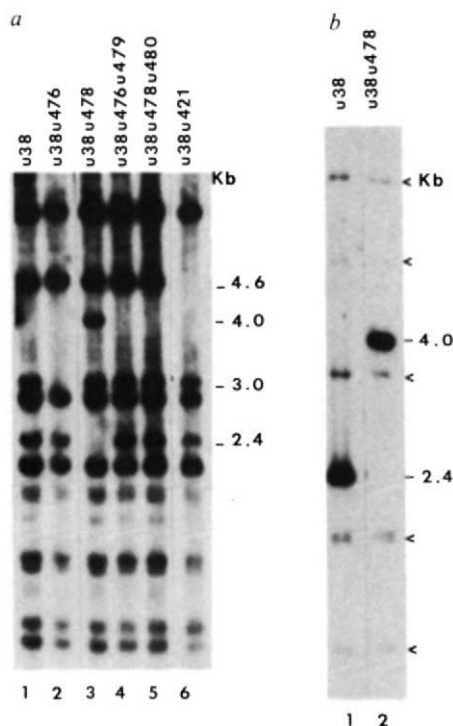


FIG. 5 Southern blot analysis of *deg-1* DNAs. *a*, Genomic DNA digested with *EcoRI* and *XbaI* and probed with cosmid C47C12. Lane 1, *deg-1(u38)*; lane 2, a *mut-2*-derived revertant of *deg-1(u38)* containing a Tc1 insertion in the 3.0 kb *EcoRI*-*XbaI* fragment (fragment I in Fig. 6); the resulting fragment with the insertion comigrates with a second fragment and increases the intensity of the band at 4.6 kb; lane 3, a second revertant with an insert in the 2.4 *EcoRI* fragment (fragment IC in Fig. 6); lanes 4 and 5, excision strains from *deg-1(u38u476)* and *deg-1(u38u478)*, respectively, which express the Deg and Tab phenotypes; lane 6, an EMS-derived revertant missing the 4.5 kb *EcoRI* fragment (fragment D in Fig. 6). No differences in the hybridization pattern were detected between *deg-1(u38)* and wild-type DNA probed with cosmid C47C12 or R02A8 (see Fig. 6). A third *mut-2*-derived revertant, *u38u477*, also had the same pattern of hybridization as the wild type. *b*, Genomic digested with *EcoRI* and *XbaI* and probed with fragment IC. The IC fragment hybridizes to the *deg-1* cDNA and contains the insertion site of the Tc1 transposon in revertant strain *u38u478*. Lane 1, *deg-1(u38)*; lane 2, *deg-1(u38u478)*. Five cross-hybridizing bands, which are also seen in the wild type in other blots, are indicated by arrowheads.

METHODS. *C. elegans* genomic DNA was prepared as described⁶¹. General molecular methods were as described⁶², with final washes of blots in $0.1 \times \text{SSC}$ at 65°C . Initially *deg-1* DNA was cloned from the *u38u476* strain. Southern blots of DNA from this strain, when probed with DNA from the transposon Tc1⁶¹, revealed a single, novel Tc1-containing *EcoRI*-*XbaI* fragment that cosegregated with the non-Tab phenotype in recombinants with *u38*. This fragment was isolated from a plasmid library constructed of size-fractionated *EcoRI*-*XbaI* fragments from *u38u476* DNA in plasmid pUC18. The resulting plasmid was digested with *EcoRV* to remove the Tc1 element and religated to produce plasmid TU#8.

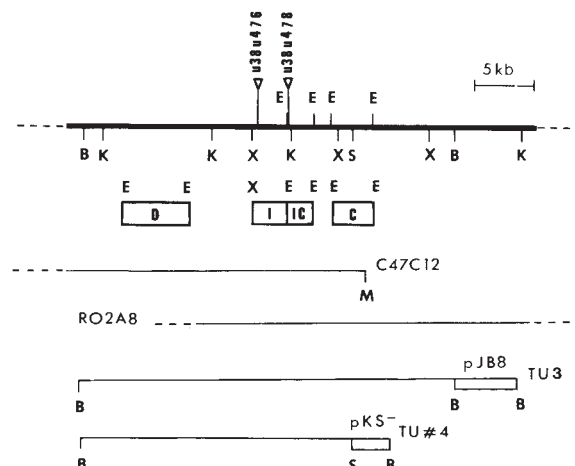


FIG. 6 DNAs of the *deg-1* region. The top line represents a partial restriction map of genomic DNA. Triangles denote the approximate positions of transposon insertions in the indicated revertants (all insertions are 1.6 kb, the size of the transposon Tc1, but the presence of this transposon has been confirmed only in strain *u38u476*). The boxes under the restriction map indicate fragments possessing DNA rearrangements in *u38* revertants or that hybridize to *deg-1* cDNAs: D is deleted in *u38u421*; I contains the Tc1 insertion in *u38u476*; IC contains the transposon insertion in *u38u478* and hybridizes to the *deg-1* cDNAs (the transposon resides in a 187 bp *EcoRI*-*NdeI* fragment at the 5' end of IC; the sequence of this fragment is identical with the *deg-1* cDNAs, except for a 50 bp intron 160 bp from the *EcoRI* site); and C is a fragment that also hybridizes to the cDNAs. C47C12 and R02A8 are cosmids containing wild-type DNA²⁴. TU#3 is a cosmid clone of the *BamHI* fragment from *deg-1(u38)*; when transformed into wild-type animals, TU#3 produces the Deg and Tab phenotypes. TU#4 is a derivative of TU#3 that lacks the 3' *Smal*-*BamHI* fragment and does not generate the Deg or Tab phenotypes on transformation into wild type. B, *BamHI*; E, *EcoRI*; K, *KpnI*; S, *SmaI*; X, *XbaI*; M, *MboI*; not all *EcoRI*, *XbaI* and *MboI* sites are indicated.

METHODS. Genomic subclones of wild-type DNA from C47C12 were made by ligation into pUC18. The TU#3 cosmid was obtained by isolating large fragments from *BamHI*-digested *deg-1(u38)* DNA from a sucrose gradient⁶² and ligating them to the *BamHI*-*HindIII* and *SalI*-*BamHI* fragments of the cosmid vector pJB8⁶³. This library was screened with all four genomic fragments (D, I, IC and C) individually as probes. Animals were transformed with TU#3 DNA by co-injection of the DNA into wild-type oocytes⁶⁴ with either pPD10.41 or pGB3.5, plasmids that contain an antisense construct of the *C. elegans unc-22* gene that produces a twitching phenotype (A. Fire, personal communication). About half of the twitcher transformants also segregated animals that were Tab and Deg. *uln1* resulted from co-injection of TU#3 and pGB3.5 and showed linkage with *unc-18(e81)* on the X chromosome. Transformations with TU#4 were done similarly; Southern blotting confirmed that TU#4 was present in the transformed animals.

1	10	20	30	40	50	60	550	560	570	580	590	600
GA ATT CGT GTC CTT CTA TTC GTA AAC ACA TCA GAT TAT ATG TCA ACT TCT GAG TCA TCC GGA							GGT ATG ACA GAA AGC GAG TGC GAA CAA TAC TAT CGG CTA AAT GCG GCA ATG ATC GAG					GTA
Ile Arg Val Leu Leu Phe Val <u>Asn Thr Ser</u> Asp Tyr Met Ser Thr Ser Glu Ser Ser Gly							Gly Met Thr Glu Ser Glu Cys Glu Gln Tyr Tyr Arg Leu Asn Ala Ala Met Ile Glu					Val
1							181	*	190			200
70	80	90	100	110	120							
GTT CGA CTG GCC ATC CAT CCA CCA ACT GAG TAC CCG TTC CCC GAC ACA TTC GGC TAT TCT							610	620	630	640	650	660
Val Arg Leu Ala Ile His Pro Pro Thr Glu Tyr Pro Phe Pro Asp Thr Phe Gly Tyr Ser							Phe Tyr Glu Gln Leu Asn Tyr Glu Leu Leu Gln Glu Ser Glu <u>Ala Tyr Gly Leu Val Asn</u>					
21							201	210				220
130	140	150	160	170	180							
GCG CCA GTT GGT TTT GCA AGT AGT TTT GGA ATC AAA AAG AAA GTG ATG CAA AGG TTG CCA							670	680	690	700	710	720
Ala Pro Val Gly Phe Ala Ser Ser Phe Glu Ile Lys Lys Lys Val Met Gln Arg Leu Pro							<u>Leu Ile Ala Asp Phe Gly Gly His Leu Gly Leu Trp Leu Gly Phe Ser Val Ile Thr Val</u>					
41							221	230				240
190	200	210	220	230	240							
GCA CCA TAT GGA GAA TGT GTA GAA ACG AAG AAA GTT GTA GAC AGA AAT TAT ATT TAC GCG							730	740	750	760	770	780
Ala Pro Tyr Gly Glu Cys Val Glu Thr Lys Lys Val Val Asp Arg Asn Tyr Ile Tyr Ala							<u>Met Glu Val Cys Val Leu Leu Val Asp Met Ile Ser Leu Phe Phe</u> Lys Ser Arg His Glu					
61							241	*	250			260
250	260	270	280	290	300							
GGG TAC GAT TAT CAT CCA GAA GGT TGT CAT AGA AGT TGC TTC CAA AAT GGA CTG ATT GAT							790	800	810	820	830	840
Gly Tyr Asp Tyr His Pro Glu Gly Cys His Arg Ser Cys Phe Gln Asn Gly Leu Ile Asp							GAA AAA CTT CTG AGA CAG AGC ACA AAA AGG AAA GAT GTT CCA GAA GAT AAA CGG CAA ATT					
81							261	270				280
310	320	330	340	350	360							
GAT TGT TCG TGT GGA GAT CCT CGT TTC CCA GTA CCA GAA GGT TAT AGA CAT TGC TCG GCA							850	860	870	880	890	900
Asp Cys Ser Cys Gly Asp Pro Arg Phe Pro Val Pro Glu Gly Tyr Arg His Cys Ser Ala							ACA GTT GGA TCA GGG CGA AAG TCA GAC GCT TTC GTA TCA ATA TAA ATA CCA ACT CTC TTT					
101	*	*	90	*	110	120	281	290				300
370	380	390	400	410	420							
TTT AAT GCA ACA GCT CGT ACC TGT CTT GAG AAG AAC ATT GGC TCA GTT GGA GAT TTC CAT							910	920	930	940	950	960
Phe <u>Asn Ala Thr</u> Ala Arg Thr Cys Leu Lys Asn Ile Gly Ser Val Gly Asp Phe His							TGA ACA ACT ATT ATA TAA CGT TAA TTT TGA ACT GGG TTT CTC AAG ATG TAG TAT ACA ATG					
121							970	980	990	1000	1010	1020
430	440	450	460	470	480							
CAT ATC ACT CAA AAA ATG GAC AAA TGC GTG TGT AAG CAA TCA TGT GAA GAA ATT ATT CAT							CTG TAA CAC GTT TCA CCT TCA TTC GTT TTT TCC GAT CTC TAA TTG TAT ATA GTG AGC TTT					
His Ile Thr Gln Lys Met Asp Lys Cys Val Cys Lys Gln Ser Cys Glu Glu Ile Ile His							110	120				130
141							141	150				160
490	500	510	520	530	540							
GAA GTT ACC TTT TCA TGC TCC AAA TGG CCT TCG GGA GCT ACT GAC CTT GGA GAC TGT GAT							110	120				130
Glu Val Thr Phe Ser Cys Ser Lys Trp Pro Ser Gly Ala Thr Asp Leu Gly Asp Cys Asp							161	170				180
161												

FIG. 7 DNA and predicted protein sequence of *deg-1* cDNA. The sequence of the larger cDNA is shown. The smaller cDNA lacks the DNA encoding residues 176–199 (bracketed). Potential glycosylation sites²⁶ are underlined, cysteine residues are asterisked, and the hydrophobic region is double-underlined.

METHODS. The cDNAs were subcloned from λ gt10 into the *Eco*RI site of

protein forms channels or modifies existing channels. Alternatively, *deg-1* may encode a ligand-activated receptor. The cysteine-rich extracellular domain suggests such a function. The *u38* phenotype could then arise from an inappropriate activation of a second messenger pathway that alters the activity of membrane proteins. A further possibility is that the *deg-1(u38)* product interferes with essential cell functions, as hypothesized for the abnormal catabolism of the β -amyloid precursor protein in Alzheimer's disease⁴².

The suppression of deaths by *mec-6* mutations indicates that an important role for *mec-6* in neurodegeneration. This *mec-6* activity could be required for the *mec-4(d)* and *deg-1* degenerations either because it activates the products of these genes (perhaps by regulating their synthesis or subcellular localization) or because it is a necessary target (such as a channel) or cofactor for their action. In all these cases, loss of *mec-6* activity would prevent the neuronal degeneration. The further molecular analysis of the *deg-1*, *mec-4* and *mec-6* genes as well as the localization of their products will be important in elucidating the ways these genes act.

The selective degeneration of particular sets of neurons is characteristic of a number of human genetic diseases¹, but the relationship between mutationally induced neurodegeneration in *C. elegans* and in humans or other organisms is unknown. Nonetheless, parallels are seen. Many late-onset human neurodegenerative diseases are expressed as dominant traits¹, indicating that they could result from the production of abnormal products. Although more dominant human genetic diseases are known than recessive or X-linked traits⁴³, the proportion of dominant, late-onset diseases of the nervous system seems particularly high. Another similarity between the *C. elegans* degenerations and those in human diseases may be their appearance. It is difficult to determine whether the pathologies

pKS⁻ and pKS⁺ (Stratagene). Single stranded DNA was rescued from a set of nested deletions⁶⁵ using helper phage R408 (Stratagene) and sequenced by the dideoxy termination method⁶⁶ using either Klenow (NEB) or Sequenase (USB). Both strands were sequenced. Sequence comparison to Genbank, PIR and Clavier DNA and protein databases was performed using the DFASTN and DFASTP.

described for the human diseases are identical to the vacuolated deaths seen in the *C. elegans* mutants. Yet in humans, vacuolated cortical cells ('ballooned neurons') occur in some patients with neurodegenerative diseases, including instances of Alzheimer's disease⁴⁴, and vacuolated cells have been reported in two cases of motor neuron disease^{45,46}. Vacuolization is also seen in genetically induced neuronal death in several mouse mutants^{47–49}.

Although the *mec-4(d)* and *deg-1*-induced degenerations are distinct from programmed cell deaths in *C. elegans*, mechanisms similar to those occurring in these mutants may function not only in neurodegenerative diseases, but also in cell deaths that arise during normal development in other organisms. Several researchers have reported that some, but not all naturally occurring neuronal deaths in vertebrates are characterized by an initial dilation and vacuolization of the cytoplasm of the affected cells^{50–53}. The process of cell death in these instances may be similar to the *C. elegans* degenerations. Such cell death could, as with the PVC cells in *deg-1(u38)* mutants, exhibit delayed onset. Such a process may occur in the death of the subplate neurons of the mammalian cerebral cortex which die with a vacuolated appearance after making functional embryonic synapses^{54,55}. It seems likely that cell death during normal development occurs through more than one mechanism, one of which could utilize a *deg-1(u38)*- or *mec-4(d)*-like product. □

Received 20 November 1989; accepted 26 March 1990.

- Adams, R. D. & Victor, M. in *Principles of Neurology* 3rd edn Ch. 42 859–901 (McGraw-Hill, New York, 1985).
- Hayden, M. R. *Huntington's Chorea* (Springer, Berlin, 1981).
- Gusella, J. F. et al. *Nature* **306**, 234–238 (1983).
- Mulder, D. W., Kurland, L. T., Offord, K. P. & Beard, M. *Neurology* **36**, 511–517 (1986).
- Davies, P. *Neurobiol. Aging* **7**, 459–466 (1986).
- Harding, A. E. in *Hereditary Ataxias and Related Disorders* Ch. 10, 129–165 (Churchill Livingstone, Edinburgh, 1984).

7. Chalfie, M. & Au, M. *Science* **243**, 1027-1033 (1989).
8. Chalfie, M. & Sulston, J. *Dev Biol* **82**, 358-370 (1981).
9. Chalfie, M. et al. *J. Neurosci.* **5**, 956-964 (1985).
10. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. *Dev Biol* **100**, 64-119 (1983).
11. Sulston, J. E. & Horvitz, H. R. *Dev Biol* **56**, 110-156 (1977).
12. Herman, R. K. *Genetics* **116**, 377-388 (1987).
13. Way, J. C. & Chalfie, M. *Genes Dev.* **3**, 1823-1833 (1989).
14. Brenner, S. *Genetics* **77**, 71-94 (1974).
15. Greenwald, I. S. & Horvitz, H. R. *Genetics* **96**, 147-164 (1980).
16. Collins, J., Saari, B. & Anderson, P. *Nature* **328**, 726-728 (1987).
17. Suzuki, D. T. *Science* **170**, 695-706 (1970).
18. Ferguson, E. L. & Horvitz, H. R. *Genetics* **123**, 109-121 (1989).
19. Ellis, H. M. & Horvitz, H. R. *Cell* **44**, 817-829 (1986).
20. Avery, L. & Horvitz, H. R. *Cell* **51**, 1071-1078 (1987).
21. Choi, D. *Neuron* **1**, 623-634 (1988).
22. White, J. G., Albertson, D. G. & Anness, M. A. *Nature* **271**, 764-766 (1978).
23. Sulston, J. E. & Horvitz, H. R. *Dev Biol* **82**, 41-55 (1981).
24. Coulson, A., Sulston, J., Brenner, S. & Karn, J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7821-7825 (1986).
25. Savage, C. et al. *Genes Dev.* **3**, 870-881 (1989).
26. Hubbard, S. C. & Ivatt, R. J. A. *Rev. Biochem.* **50**, 555-583 (1981).
27. Adams, G. A. & Rose, J. K. *Cell* **41**, 1007-1015 (1985).
28. Yanagi, Y. et al. *Nature* **308**, 146-149 (1984).
29. Hedrick, S. M., Nielsen, E. A., Kavalier, J., Cohen, D. I. & Davis, M. M. *Nature* **308**, 153-158 (1984).
30. Yamamoto, T. et al. *Cell* **39**, 27-38 (1984).
31. Ullrich, A. et al. *Nature* **309**, 418-425 (1984).
32. Wen, D. et al. *Biochemistry* **26**, 4350-4357 (1987).
33. Radeke, M. J., Misko, T. P., Hsu, C., Herzenberg, L. A. & Shooter, E. M. *Nature* **325**, 593-597 (1987).
34. Trent, C., Tsung, N. & Horvitz, H. R. *Genetics* **104**, 619-647 (1983).
35. Ferguson, E. L. & Horvitz, H. R. *Genetics* **110**, 17-72 (1985).
36. Folstein, S. E. et al. *Science* **229**, 776-779 (1985).
37. Leonard, J. P. & Salpeter, M. M. *J. Cell Biol.* **82**, 811-819 (1979).
38. Hogan, P. G. *J. Neurosci.* (in the press).
39. Choi, D. W. *J. Neurosci.* **7**, 369-379 (1988).
40. Takumi, T., Ohkubo, H. & Nakanishi, S. *Science* **242**, 1042-1045 (1988).
41. Murai, T., Kakizuka, A., Takumi, T., Ohkubo, H. & Nakanishi, S. *Biochem. biophys. Res. Commun.* **161**, 176-181 (1989).
42. Kang, J., Lemaire, H.-G., Unterbeck, A., Salbaum, J. M. & Masters, C. L. *Nature* **325**, 733-736 (1987).
43. McCusick, V. A. *Mendelian Inheritance in Man* 8th edn (Johns Hopkins University Press, Baltimore, 1989).
44. Dickson, D. W. et al. *Acta neuropath.* **71**, 216-223 (1986).
45. Kohn, R. *J. Neurol. Neurosurg. Psychiat.* **34**, 427-431 (1971).
46. Reif-Kohn, R. & Mundel, G. *Confin. Neurol.* **36**, 23-32 (1974).
47. Duchen, L. W. & Strich, S. J. *J. Neurol. Neurosurg. Psychiat.* **31**, 535-542 (1968).
48. Caddy, K. W. T. & Briscoe, T. J. *Brain Res.* **91**, 276-280 (1975).
49. LaVail, M. M., Blanks, J. C. & Mullen, R. J. *J. comp. Neurol.* **212**, 217-230 (1982).
50. Pilar, G. & Landmesser, L. *J. Cell Biol.* **68**, 339-356 (1976).
51. Chu-Wang, I.-W. & Oppenheim, R. W. *J. comp. Neurol.* **177**, 33-58 (1978).
52. Knyihar, E., Csillik, B. & Rakic, P. *Science* **202**, 1206-1209 (1978).
53. Cunningham, T. J. *Int. Rev. Cytol.* **74**, 163-186 (1982).
54. Valverde, F. & Facal-Valverde, M. V. *J. comp. Neurol.* **269**, 168-192 (1988).
55. Chun, J. J. M. & Shatz, C. J. *J. comp. Neurol.* **282**, 555-569 (1989).
56. Way, J. C. & Chalfie, M. *Cell* **54**, 5-16 (1988).
57. Hedgecock, E. M., Sulston, J. E. & Thomson, J. N. *Science* **220**, 1277-1279 (1983).
58. Chalfie, M. & Thomson, J. N. *J. Cell Biol.* **82**, 278-289 (1979).
59. Ward, S., Thomson, N., White, J. G. & Brenner, S. *J. comp. Neurol.* **160**, 313-337 (1975).
60. White, J. G., Southgate, E., Thomson, J. N. & Brenner, S. *Phil. Trans. R. Soc. B314*, 1-340 (1986).
61. Emmons, S. C., Klass, M. R. & Hirsh, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1333-1337 (1979).
62. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
63. Ish-Horowitz, D. & Burke, J. F. *Nucleic Acids Res.* **9**, 2989-2998 (1981).
64. Fire, A. *EMBO J.* **5**, 2673-2680 (1986).
65. Ozkaynak, E. & Putney, S. D. *BioTechniques* **5**, 770-773 (1987).
66. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).

ACKNOWLEDGEMENTS. This paper is dedicated to the memory of Helen Bass, Oscar Horvitz and Sumiko Okamoto. We thank Connie Masuoka, Lowell Chapnick, Cathy Savage, Michael Drucker and Jennifer Mach for help in these experiments; Jim Thomas for sharing data on the AVA ablations; and Bob Horvitz, Bob Brown, Pat Hogan, Carla Shatz, Anne Chambers, Monica Driscoll, Chris Li and Jeff Way for discussion. E.W. was a postdoctoral fellow of the Jane Coffin Childs Memorial Fund for Cancer Research and of the Huntington's Disease Society of America. This work was supported by the US Public Health Service, the Muscular Dystrophy Association of America and Cambridge Neuroscience Research, Inc.

LETTERS TO NATURE

Rapid recovery of the Vela pulsar from a giant glitch

Claire S. Flanagan

Hartebeesthoek Radio Astronomy Observatory, PO Box 443, Krugersdorp 1740, South Africa

OBSERVATIONS of the rotational behaviour of pulsars have shed much light on the structure of neutron stars^{1,2}. In particular, the response of the Vela pulsar (PSR0833-45) following each of seven giant glitches (sudden increases in rotation rate ν of magnitude $\Delta\nu/\nu \approx 10^{-6}$, and in the derivative, $\Delta\dot{\nu}/\dot{\nu} \approx 10^{-2}$) has been a recovery on two timescales, of about 6 and 60 days. This has been interpreted as evidence for the existence of distinct regions of neutron superfluid within the stellar crust². Here I report observations that started 35 minutes after the eighth glitch of PSR0833-45, which reveal a hitherto unseen glitch component: a $(12 \pm 1)\%$ increase in the slowdown rate $\dot{\nu}$ (the largest so far seen), most of which decayed away over 0.4 days. These data support the existence of at least two superfluid components in the stellar crust that are coupled linearly to the normal crust, as suggested recently by Alpar et al.³.

The Vela pulsar has been monitored nearly daily since 1984 at 1.644 and 2.3256 GHz, with the 26-m antenna at the Hartebeesthoek Radio Astronomy Observatory. During 1988, observations were made hourly to increase the chances of observing a glitch. Shortly after the pulsar rose on 24 December 1988 (JD 2447520) it spun up^{4,5}. This giant glitch was detected at the next observation 35 minutes later. Continuous data capture was automatically initiated, and continued for the next 10 days, for the $11\frac{1}{2}$ hours per day that the pulsar is observable. Less frequent observations have been made since then.

Pulse arrival times were obtained from 500-pulse (45-s) on-line integrations of the pulsed signal, synchronized with the apparent pulse period, using conventional equipment and reduction

methods. Time was provided by a hydrogen maser, and referred to universal time using the global positioning satellite system. Station arrival times were reduced to the Solar System barycentre and to infinite observing frequency using the Harvard-Smithsonian Center for Astrophysics PEP 740R ephemeris and reduction software (J. F. Chandler, personal communication). The optical position⁶, with zero proper motion, was used. The resulting averaged arrival times had a short-term r.m.s. scatter of 10-20 μ s.

The pre-glitch rotation parameters (rotation frequency ν and

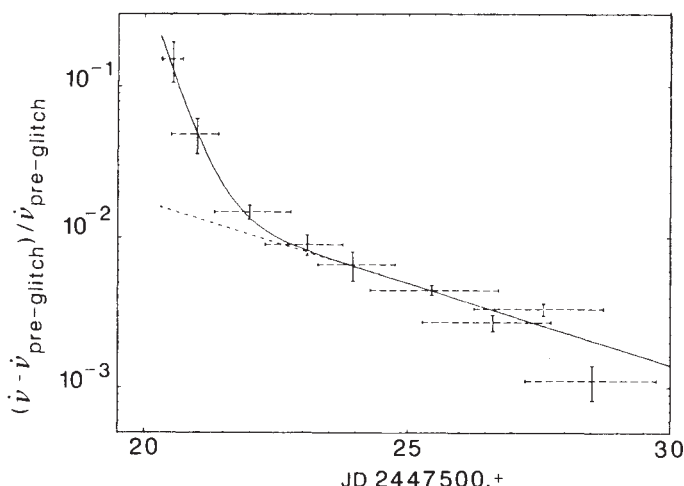


FIG. 1 Residuals in $\dot{\nu}$ for PSR0833-45 after removal of the linear and longest exponential decay terms listed in Table 1. The dashed curve is the 4-d decay from Table 1; the solid curve is the sum of the 4-d and 0.4-d decays. The dotted vertical line marks the glitch epoch. Data are normalized to the pre-glitch value of $\dot{\nu}$. Vertical error bars are $\pm 1\sigma$. Horizontal error bars show the range of data fitted.

TABLE 1 PSR0833–45 rotation parameters at the time of the 1988 glitch

Pre-glitch parameters	
ν_0	11.20001539350(2) Hz
$\dot{\nu}_0$	$155.8232(4) \times 10^{-13} \text{ Hz s}^{-1}$
$\ddot{\nu}_0$	$8.3(9) \text{ Hz s}^{-2}$
Epoch, t_0	JD 2447493.15
Range fitted	JD 2447480–2447520
RMS residual	180 μs
Glitch parameters	
$\Delta\nu/\nu$	$1.8071(8) \times 10^{-6}$
$\Delta\dot{\nu}/\dot{\nu}$	$12(1) \times 10^{-2}$
Epoch	JD 2447520.303(1)
Range fitted	JD 2447480–2447521.76
Post-glitch decay	
$\Delta\dot{\nu}_1/\dot{\nu}$	$20(9) \times 10^{-2}$
$\Delta\dot{\nu}_2/\dot{\nu}$	$1.6(3) \times 10^{-2}$
$\Delta\dot{\nu}_3/\dot{\nu}$	$0.291(6) \times 10^{-2}$
τ_1	0.4(1) d
τ_2	4.0(5) d
τ_3	96(5) d
$(\dot{\nu}_0 - \dot{\nu}_{\text{pre-glitch}})/\dot{\nu}$	$0.488(4) \times 10^{-2}$
$\ddot{\nu}_0$	$8 \times 10^{-22} \text{ Hz s}^{-2}$ (fixed)
Range fitted	JD 2447520.3–2447840

The numbers in parentheses are the 1σ errors quoted in the last digit.

its derivatives) and their changes at the time of the glitch ($\Delta\nu/\nu$, $\Delta\dot{\nu}/\dot{\nu}$) are listed in Table 1. The change in frequency derivative $\dot{\nu}$ decays rapidly following the glitch, so $\Delta\dot{\nu}/\dot{\nu} = (12 \pm 1)\%$, obtained from a fit to 1.5 days of post-glitch data that assumed constant $\ddot{\nu}$, must be regarded as a lower limit. This value of $\Delta\dot{\nu}/\dot{\nu}$ is three times any seen before in PSR0833–45 and twice the previous largest value, seen in PSR0355+54 (ref. 7). The glitch epoch agrees with that reported by Hamilton *et al.*⁵, who were observing the pulsar at the time of the event.

To study the post-glitch behaviour of the pulsar, models for the recovery of $\dot{\nu}$ were fitted to 320 days of post-glitch data. Successive short spans of data were fitted for ν and $\dot{\nu}$, resulting in a set of $(\dot{\nu}_i, t_i)$ values. These data were then fitted with a sum of simple exponential decays and a linear decrease in $|\dot{\nu}|$ following the glitch epoch t_g

$$\dot{\nu}(t) = \dot{\nu}_0 + \ddot{\nu}_0(t - t_g) + \sum_i \Delta\dot{\nu}_i \exp[-(t - t_g)/\tau_i]$$

This method enhances the detectability of the decay in $\dot{\nu}$ above the 'timing noise', which is particularly evident in the Vela pulsar as large systematic excursions in rotation phase^{8,9}.

Cordes *et al.*⁸ obtained good fits to data following six glitches by using two exponential decays, of characteristic timescales $\tau_i \approx 6$ d and $\tau_i \approx 60$ d. The 1988 glitch is best fitted with two decays with similar parameters, plus another, much shorter (0.4 ± 0.1 d), exponential term which contains most of the jump in $\dot{\nu}$ ($\Delta\dot{\nu}/\dot{\nu} = 0.20 \pm 0.09$) (Table 1, Fig. 1). As the $(\dot{\nu}_i, t_i)$ data were obtained from fits to ≥ 12 hours of arrival times, the 0.4-d timescale is an upper limit. Because of the relatively short span of data, various combinations of the longest decay and linear term ($\ddot{\nu}$) could be fitted with equal success. An average value of $8 \times 10^{-22} \text{ Hz s}^{-2}$ was assumed for $\ddot{\nu}$ to obtain the parameters listed in Table 1.

In the 'vortex creep' model which Alpar *et al.* developed² and tested on the first four Vela glitches¹⁰, and on glitches in the Crab pulsar and PSR0525+21 (ref. 11), the neutron star has two components: (1) The tightly coupled interior, crust and magnetosphere which constitute the bulk of the star. We observe the rotation of this component through radio emission from the magnetosphere. (2) A layer of superfluid within the crust which is relatively loosely coupled to the crust, and therefore lags the general slowdown of the star. In this model, the glitch is caused by a sudden transfer of angular momentum from part of the

crustal superfluid to the rest of the star. The rotational coupling between the crustal superfluid and crust is dependent on the difference ω between the crustal and superfluid rotation rates. The superfluid thus decouples from the star during a glitch, and recouples as ω redevelops following the glitch. This is observed as an increase in $|\dot{\nu}|$ of the crust at the glitch, followed by its decay towards the pre-glitch value. An abnormally high value of $\ddot{\nu}$ is therefore also evident during the years of post-glitch recovery, and seems to persist until the next glitch.

Alpar *et al.*² assumed that the crustal superfluid coupling depends exponentially on ω , and thus proposed that the recovery of $\dot{\nu}$ will be delayed by $t_0 = \Delta\nu/|\dot{\nu}|$, typically two to three weeks for the Vela pulsar. Because the epochs of the early Vela glitches were uncertain by one to three weeks, this prediction could not be fully tested. Signs of immediate post-glitch recoupling on timescales of 1.6 and 3.2 d were observed in two recent Vela glitches^{12,13}. The first post-glitch observations of these events were 0.8 and 0.7 d after the glitches respectively, which precluded the detection of faster decays. Similar behaviour has recently been observed in PSR0355+54 following a spin-up⁷ that had a predicted recovery delay t_0 of five years and yet recovered partially on a timescale of 44 days. This last event has been modelled³ by including a crustal superfluid component in which the superfluid-to-crust coupling depends linearly on ω . The observations reported here of two short decays immediately following the glitch imply that at least two similar components may exist in the Vela pulsar.

The pre-glitch $\dot{\nu}$ was $-(155.804 \pm 0.003) \times 10^{-13} \text{ Hz s}^{-1}$, which is in the range $-(156.0 \pm 0.3) \times 10^{-13} \text{ Hz s}^{-1}$ noted by Cordes *et al.*⁸ for the first six Vela glitches. No sign of any glitch precursor was evident in the timing data. Over the six months before the glitch, $\ddot{\nu}$ had decreased to $(1.7 \pm 0.2) \times 10^{-22} \text{ Hz s}^{-2}$. This is well below any value reported by Cordes *et al.*, and greater than the theoretical 'unglitched' value $\ddot{\nu} = n\dot{\nu}^2\nu^{-1} = 0.7 \times 10^{-22} \text{ Hz s}^{-2}$ with braking index $n = 3$. This implies that the recovery following the previous glitch may have been nearing completion, although $\ddot{\nu}$ over such short data spans varies markedly during interglitch epochs, owing to timing noise in $\dot{\nu}$.

Timing noise was measured as the r.m.s. phase residual to a 100-d fit, after removal of the three decay components. The level of timing noise following this glitch is $(60 \pm 35) \mu\text{s}$, similar to that reported by Cordes *et al.*⁸ following previous glitches. Immediately after the glitch, timing noise is higher than average ($\sim 180 \mu\text{s}$), implying that the three-component decay model may be too simple.

Following a suggestion (C. R. Gwinn, personal communication) that the disturbed star may nutate immediately following a glitch, periodic variations were searched for in the pulse shape, pulse strength and timing data of the first day following the glitch. Periodograms were constructed and evaluated according to Scargle¹⁴. No significant periodicities were found down to 2.6 min. The variations in timing and flux data of longer than a day were consistent with those expected from timing noise and interstellar scintillation respectively, and were not truly periodic. $\square$

Received 17 January; accepted 10 April 1990.

1. Baym, G., Pethick, C., Pines, D. & Ruderman, M. *Nature* **224**, 872–874 (1969).
2. Alpar, M. A., Anderson, P. W., Pines, D. & Shaham, J. *Astrophys. J.* **276**, 325–334 (1984).
3. Alpar, M. A., Cheng, K. S., Pines, D. & Shaham, J. *Mon. Not. R. astr. Soc.* **233**, 25–31 (1988).
4. Flanagan, C. *IAU Circ. No.* 4695 (1988).
5. Hamilton, P. A., King, E. A., McConnell, D. & McCulloch, P. M. *IAU Circ. No.* 4708 (1989).
6. Manchester, R. N. *et al. Mon. Not. R. astr. Soc.* **184**, 159–170 (1978).
7. Lyne, A. G. *Nature* **326**, 569–571 (1987).
8. Cordes, J. M., Downs, G. S. & Krause-Polstorff, J. *Astrophys. J.* **330**, 847–869 (1988).
9. Downs, G. S. *Astrophys. J.* **249**, 687–697 (1981).
10. Alpar, M. A., Anderson, P. W., Pines, D. & Shaham, J. *Astrophys. J.* **278**, 791–805 (1984).
11. Alpar, M. A., Nandkumar, R. & Pines, D. *Astrophys. J.* **288**, 191–195 (1985).
12. McCulloch, P. M., Hamilton, P. A., Royle, G. W. R. & Manchester, R. N. *Nature* **302**, 319–321 (1983).
13. McCulloch, P. M., Klekociuk, A. R., Hamilton, P. A. & Royle, G. W. R. *Aust. J. Phys.* **40**, 725–730 (1987).
14. Scargle, J. D. *Astrophys. J.* **263**, 835–853 (1982).

ACKNOWLEDGEMENTS. I thank G. D. Nicholson and M. J. Gaylard for comments.

Measurement of the London moment in two high-temperature superconductors

A. A. Verheijen, J. M. van Ruitenbeek,
R. de Bruyn Ouboter & L. J. de Jongh

Kamerlingh Onnes Laboratorium, Rijksuniversiteit Leiden, Postbus 9506,
2300 RA Leiden, The Netherlands

THE well-known London equations dictate that the magnetic field inside a stationary superconductor vanishes. The same equations, however, predict that the superconductor generates a magnetic field when it is put into uniform rotation. The associated magnetic moment is known as the London moment. Here we show that high-transition-temperature (high- T_c) superconducting oxides indeed possess this fundamental property, and when rotated, develop a London moment of the same magnitude as that of conventional superconductors, provided that the signal is scaled to the shielded volume fraction. This confirms that high- T_c superconductivity corresponds to a superfluid phase in the London sense, and that the mass of the carriers that appears in the expression of the London moment is the free-electron mass, irrespective of the nature of the carriers in the normal state.

The London field is given by¹

$$\mathbf{B} = -\frac{2m}{e}\boldsymbol{\omega}, \quad (1)$$

where $\boldsymbol{\omega}$ is the angular frequency, $-e$ is the electron charge and m is the mass of the charge carriers to be discussed below. The origin of the effect may be understood in several ways, the earliest argument being that of Becker *et al.*² Consider a charged superfluid in a matrix of positively charged ions, with no frictional coupling between these two systems. Consequently, when the superconductor is put into rotation one could initially view the superfluid as unaffected by this motion and remaining immobile. But the rotating positively charged ion matrix constitutes a current that generates a magnetic field. This generates an electric field by induction, so that the superfluid is set into rotation until the lattice and the superfluid move in unison, except for a surface layer of thickness λ where the superfluid electrons lag behind. The net effect is thus a surface current that

generates the magnetic field in equation (1). Alternatively, one can derive the London moment using conservation of total angular momentum, as for other gyromagnetic phenomena^{1,3}. These arguments, however, are not sufficient. Superconductivity is a macroscopic quantum phenomenon and is described by a unique ground state. As a consequence the same London moment is also developed when cooling a rotating superconductor through its transition temperature T_c . To show this we recall that the London moment can be derived from the London equation $\nabla \times \mu_0 \lambda_L^2 \mathbf{j}_s = -\mathbf{B}$, in combination with the Maxwell equation $\nabla \times \mathbf{B} = \mu_0 \mathbf{j} = \mu_0 (\mathbf{j}_s - n_s e \mathbf{v}_0)$. Here $\mathbf{j}_s$ is the supercurrent density, $\lambda_L^2 = m / \mu_0 n_s e^2$ the London penetration depth, n_s the density of superfluid electrons and $\mathbf{v}_0 = \boldsymbol{\omega} \times \mathbf{r}$ the velocity of the positively charged lattice. Solving these equations for the magnetic field yields

$$\nabla^2 \mathbf{B}_{\text{eff}} = \frac{1}{\lambda_L^2} \mathbf{B}_{\text{eff}} \quad (2)$$

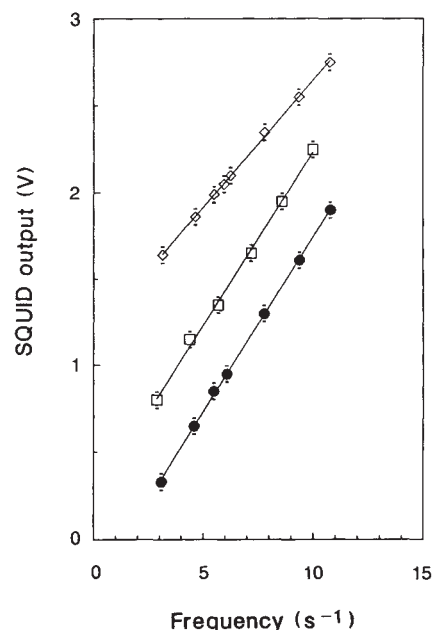
where

$$\mathbf{B}_{\text{eff}} = \mathbf{B} + 2\mu_0 n_s e \lambda_L^2 \boldsymbol{\omega} = \mathbf{B} + \frac{2m}{e} \boldsymbol{\omega} \quad (3)$$

Equations (2) and (3) tell us that except for a surface layer of thickness λ_L the field inside a rotating superconductor is equal to $-(2m/e)\boldsymbol{\omega}$, as in equation (1). Various arguments (see, for example, refs 4, 5) state that m in equation (1) should be the bare-electron mass m_e , at least if relativistic effects are disregarded⁶. In view of the debate about the nature of the carriers and of the (possibly exotic) pairing mechanisms in the high- T_c superconductors it seemed worthwhile to show that they possess the same fundamental property.

The London moment was measured on sintered $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ and $\text{BaPb}_{0.8}\text{Bi}_{0.2}\text{O}_3$. For the $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ sample the onset transition temperature as determined by resistance and magnetic measurements is 92.0 K. The width of the transition in the resistance measurement is 0.5 K. The sample shows full shielding as measured in susceptibility in an a.c. field of 24 μT . The Meissner effect determined using a SQUID (superconducting quantum-interference device) in a field of 0.2 mT corresponds to 25% of the total volume. For $\text{BaPb}_{0.8}\text{Bi}_{0.2}\text{O}_3$ the transition temperature determined in a.c. susceptibility is 10.0 K. The sample shows only 91% shielding at 4.2 K. The partial shielding is a problem generally encountered in the ceramic superconductors and may be related to inhomogeneous composition and weak coupling between grains.

FIG. 1 The change in the SQUID output voltages as a function of the frequency of rotation for $\text{BaPb}_{0.8}\text{Bi}_{0.2}\text{O}_3$ ($\diamond$), Pb ($\square$) and $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ ($\bullet$). 1 V corresponds to $(3.6 \pm 0.2) \times 10^{-10}$ T. The vertical position of the data is arbitrary, only the slope is relevant. The $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ sample shows the same London moment as the Pb sample to within the experimental error, a least-squares fit to the data points gives slopes 0.202 ± 0.007 V s and 0.198 ± 0.008 V s for $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ and Pb respectively. For the $\text{BaPb}_{0.8}\text{Bi}_{0.2}\text{O}_3$ sample the slope is 25% smaller due to non-ideal shielding and a slight difference in sample dimensions. Both samples were pressed into pellets of ~ 2 -mm thickness and 12-mm diameter and sintered at the appropriate temperatures.



The London moment was measured by rotating the disk-shaped superconductors inside a flux transformer which couples the magnetic flux to a commercial r.f. SQUID. All measurements were done by first cooling to 4.2 K and then setting the sample into rotation. This results in a signal that is proportional to the shielded volume fraction in a zero-field cooled magnetization measurement. A lead sample shaped as closely as possible to the same dimensions as the $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ sample is used to calibrate the measurement. When inserting the free-electron mass and charge in equation (1) a magnetic field of $\omega \times 1.14 \times 10^{-11}$ T is expected, with ω in rad s^{-1} . With the measured self-inductance of the flux transformer and an estimate of the fraction of the flux actually detected, the intensity of the magnetic field measured for the lead sample is within 10% of this value. The 10% error is the result of uncertainties in the geometrical factors and the actual self-inductance of the flux transformer with the superconductor in the primary coil. This error only affects the absolute values. For the measurement on the high-temperature superconductors the lead sample serves as a reference.

Figure 1 shows that the magnetic moment varies linearly with frequency. The moment changes sign when the sense of rotation is reversed. The sign of the signals of the three samples is the same. From Fig. 1 it is clear that the signals of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ and Pb are equal, within the experimental error. For $\text{BaPb}_{0.8}\text{Bi}_{0.2}\text{O}_3$ the signal is about 25% smaller than that of the Pb reference sample. This may be attributed to the non-ideal shielding found for the $\text{BaPb}_{0.8}\text{Bi}_{0.2}\text{O}_3$ sample, and to slightly different sample dimensions. When corrections are applied for the coupling factor of the signal to the flux transformer and the signal is scaled to the measured shielded volume fraction, it is again equal to the Pb reference signal to within the estimated accuracy. We conclude that the abundance of defects and grain boundaries in our sintered high- T_c materials have no other effect on the London moment than to reduce it in proportion to the reduced shielding.

From these observations we conclude that the superconductivity in $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ and $\text{BaPb}_{0.8}\text{Bi}_{0.2}\text{O}_3$ corresponds to a superfluid phase in the London sense. The value of the London moment is equal to that of conventional BCS superconductors. From this it follows in particular that the mass in equation (1) is the free-electron mass irrespective of the material. This last result is of some interest in view of the special nature of the oxide superconductors, which can be thought of as being derived from semiconductors by suitable doping or varying the oxygen concentration (self-doping). The conductivity in the normal state is associated with negatively charged carriers in $\text{BaPb}_{0.8}\text{Bi}_{0.2}\text{O}_3$ but with positive carriers (holes) in the case of $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$. Because of the low carrier densities and very small coherence lengths, it has been proposed (see, for example, refs 7, 8) that real-space pairing is the superconductivity mechanism, as opposed to the $\mathbf{k}$ -space (Cooper) pairing of the BCS theory. One might expect heavy masses for such carriers but this work demonstrates that the effective-mass concept is irrelevant for the London moment, that is, it involves the bare mass m_e irrespective of the nature of the carriers in the normal state. This may be understood from the fact that the London moment is a magneto-mechanical effect³ and therefore involves the true mechanical momentum $m_e \mathbf{v}$ of the carriers (see also refs 4, 5). □

Three-dimensional twisted vortices in an excitable chemical medium

A. M. Pertsov, R. R. Aliev & V. I. Krinsky

Institute of Biophysics, USSR Academy of Sciences, 142292 Pushchino, Moscow Region, USSR

SEVERAL different types of three-dimensional nonlinear wave structure are expected to occur in chemically excitable media¹. In general these are vortex-like structures (Fig. 1, for example) in which the waves adopt complex spatial configurations (curved, knotted, twisted or closed into rings). But until recently only simple vortices and vortex rings have been observed experimentally, for example in the Belousov-Zhabotinsky (B-Z) reaction²⁻⁴ and in cardiac tissue⁵. Here we report a new type of structure, twisted vortices, observed in a B-Z reaction immobilized in agarose gel (Fig. 2). In cross-section, a twisted vortex resembles a rotating spiral with a phase of rotation that changes along the filament. The filament extends through the entire medium, terminating at its free surfaces. The wavelength of a twisted vortex is shorter than that of a simple one and decreases with increasing twist, in accord with theory^{6,7}. The period shows little dependence on twist, contrary to the results of computer experiments⁶. Twisted vortices decay into simple ones; their twist decays exponentially with a time constant that increases with filament length. These dynamics are consistent with the theory of diffusional untwisting^{8,9} rather than that of shock waves¹⁰.

In a three-dimensional excitable medium, simple vortices are formed readily (Fig. 1a). These are rotating spiral waves that represent a straightforward extension of those seen in two dimensions, the rotation phase being independent of the third coordinate. Most of the spiral waves observed experimentally (for example, in the B-Z reaction^{2-4,11,12}, in cardiac muscle¹³, in the retina¹⁴ and in slime molds¹⁵), are such simple three-dimensional vortices. But numerical experiments with a Fitz-Hugh-Nagumo model¹⁶ and analytical theory⁶ predict that twisted vortices should have smaller periods and wavelengths than other types of autonomous waves and thus should dominate over these by suppressing and expelling them.

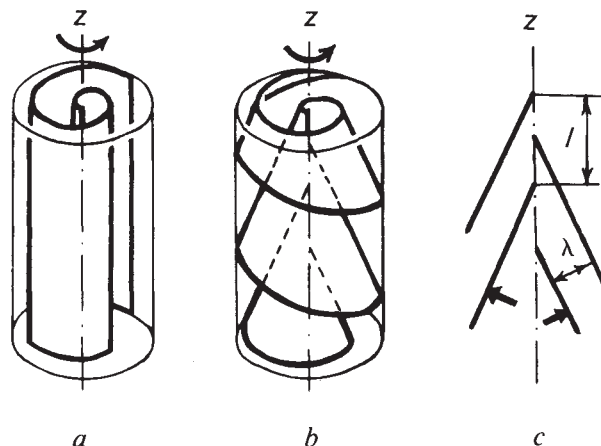


FIG. 1 Schematic representations of simple (a) and twisted (b, c) vortices. The top view of a twisted vortex is a rotating spiral wave. The side view comprises overlapping cones, seen in cross-section in c. The dash-dotted line is the rotation axis, and solid lines represent wave fronts. The directions of wave propagation and rotation are indicated by arrows. λ is the wavelength and l is the distance at which the phase shift is 2π . The twist $w = 2\pi/l$.

Received 7 March; accepted 18 April 1990.

1. London, F. *Superfluids* Vol. 1 (Wiley, New York, 1950).
2. Becker, R., Heller, G. & Sauter, F. *Z. Physik* **85**, 772-787 (1933).
3. Broer, L. J. F. *Physica* **13**, 473-478 (1947).
4. Cabrera, B., Gutfreund, H. & Little, W. A. *Phys. Rev.* **B25**, 6644-6654 (1982).
5. Cabrera, B. & Peskin, M. E. *Phys. Rev.* **B39**, 6425-6430 (1989).
6. Tate, J., Cabrera, B., Felch, S. B. & Anderson, J. T. *Phys. Rev. Lett.* **62**, 845-848 (1989).
7. Micnas, R., Ranninger, J. & Robaszkiewicz, S. *Rev. Mod. Phys.* **62**, 113-171 (1990).
8. de Jongh, L. J. *Physica* **C152**, 171-216 (1988); *Physica* **C161**, 631-655 (1989).

ACKNOWLEDGEMENTS. We thank J. M. J. van Leeuwen for his interest and for discussions. This investigation is part of the Dutch National Research Program on high-temperature superconductivity.

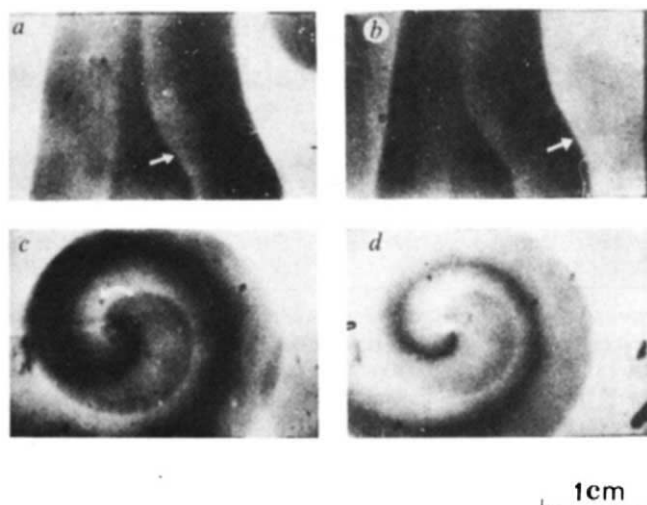


FIG. 2 A twisted vortex ($w=3\text{ cm}^{-1}$) in a chemically excitable medium. *a*, *b*, Side views (compare Fig. 1*c*); *c*, *d*, top views (compare Fig. 1*b*). The interval between *a* and *b*, and *c* and *d*, is 4 min. Arrows show the direction of wave propagation. Light is absorbed mostly by the portions of waves that lie perpendicular to the plane of the figure, so that in *a* and *b* only the central cross-section of the twisted vortex (Fig. 1*c*) is seen. The excitable medium is a B-Z reagent immobilized in agarose gel^{4,17,18}. The recipe was chosen to make photography of the thick layers possible: (50 mM NaBrO₃, 50 mM CH₂(COOH)₂, 150 mM H₂SO₄, 0.65 mM ferroin, 0.75% agarose.) The temperature is 20 °C.

We obtained twisted vortices from simple ones by creating temperature gradients along vortex filaments. Simple vortices with straight-line filaments were obtained by the standard method¹⁶. In the middle of a rectangular glass cuvette containing immobilized B-Z reagent, we placed a vertical plastic screen 0.1 mm thick, thus dividing the cuvette into two compartments. In one of these, a cylindrical wave was initiated by inserting a vertical thin silver wire. On reaching the screen, the wave broke along the generatrix (a line whose motion generates a surface) of the cylinder, the breaking edges corresponding to two straight lines which moved away from each other along the screen surface. When one of these reached the wall of the cuvette, the screen was removed and the other breaking edge curled into a simple scroll with a straight-line filament. A glass top was placed on the cuvette, so that both ends of the filament were in contact with glass, to avoid inhibition of the reaction by oxygen.

To obtain a twisted scroll the cuvette was placed on a thermostat-controlled stage heated above room temperature. This caused a local increase in the reaction rate and in the velocity of vortex rotation. An increase in temperature by 1 °C caused a decrease in the period from about 10 min to 0.5 min. These differences in local rotation velocities resulted in the development of a twist. For unequal angular frequencies of rotation, ω , the phase difference $\Delta\phi$ between two cross-sections perpendicular to the vertical (z -) axis grows linearly with time ($\Delta\phi = t\Delta\omega$), so vortices with different twists w ($w = d\phi/dz$) could be obtained by varying the time and heating temperature.

Figure 2 shows a twisted vortex. The top view corresponds to a rotating spiral wave, and the side view shows cones produced by the wave fronts. As the vortex rotates, the waves move from the centre (indicated by the arrow) to the periphery at a constant velocity.

We have studied the dependence of vortex characteristics on twist. The average twist w is defined as $w = 2\pi/l$ (see Fig. 1*c*); l was measured from the photographs. An increase in w causes a decrease in wavelength λ (Fig. 3*c*) and a weaker decrease in the rotation period (Fig. 3*b*). These are consequences of the dispersion relation for the B-Z reaction. Our measurements of the dispersion curve showed that a twist-induced change in

wavelength from 0.5 to 0.9 cm (as in Fig. 3*c*) corresponds to a slow change in rotation period from 9.0 to 10.5 min (similar to Fig. 3*b*).

When the twist exceeds a critical value w_c ($\approx 7\text{--}9\text{ cm}^{-1}$), the regular rotation of the twisted vortex is disturbed and breaks in the wave occur (arrow in Fig. 3*a*). Further increases in w give rise to a pattern resembling turbulence¹⁹. There exists a minimum wavelength for a stable wave, $\lambda_{\min}$, which can be estimated from the dispersion relation ($\lambda_{\min} \approx 0.45\text{ cm}$); λ_c is found to be close to this value.

In theory, vortices with twists below the critical value should evolve into simple ones. Spontaneous untwisting of a twisted vortex was first demonstrated in numerical experiments⁷. Analytical theories^{8–10}, however, predict different untwisting kinetics. The theory of shock waves, based on the first-order differential equation $w_t = \omega_w w_z$, states that the wave of untwisting propagates with velocity ω_w along the twisted vortex¹⁰. A more accurate equation describing the evolution of a

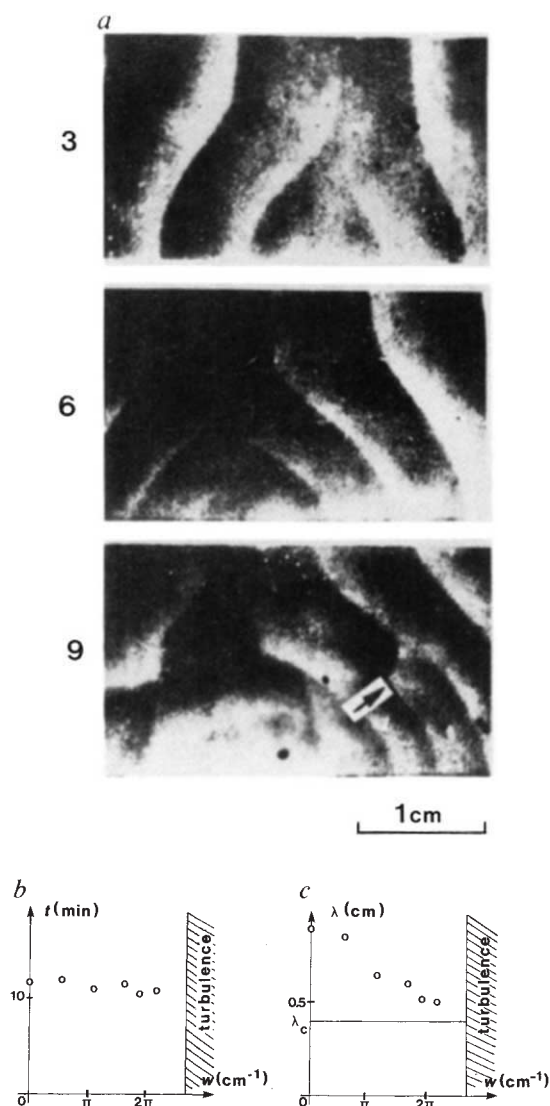


FIG. 3 Twisted vortices with different twists. *a*, Side views (compare Fig. 2*a*, *b*). Numbers on the left denote the average twist w in cm^{-1} . At $w=9\text{ cm}^{-1}$ a wave break is seen (arrow). *b*, *c*, Dependence of rotation period T and wavelength λ on twist w . T and λ were measured when the maximum temperature variation within the cuvette (measured with MT54M microthermistors) was no greater than 0.15 °C. The errors in the vortex rotation periods due to residual temperature gradients do not exceed 1%.

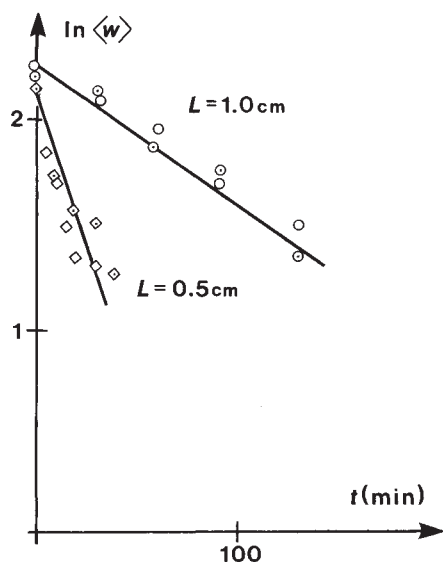


FIG. 4 Evolution of twist w . For sample thicknesses $L = 1.0$ and 0.5 cm, the kinetics of untwisting are seen to be exponential. Two sets of data are shown for each thickness. A least-squares fitting of the straight lines gives time constants τ as 161 min and 45 min respectively.

twist was proposed in ref. 8:

$$w_t = \omega_w w_z + D w_{zz} \quad (1)$$

Solution of this equation for $\omega_w = \text{constant}$ and boundary conditions $W_z|_{z=0,L} = 0$ (where L is the cuvette length) predicts exponential untwisting:

$$w(t) \propto \exp(-t/\tau); \tau = (D\pi^2/L^2 + \omega_w^2/4D)^{-1} \quad (2)$$

Our experiments confirm that vortices with twists below the critical value transform to simple ones. Moreover, the kinetics of untwisting were found to be exponential (Fig. 4) and at variance with the simple shock-wave theory of ref. 10; the time constant τ increases with L .

Equation (2) fits our experiments qualitatively but not quantitatively. Equation (2) predicts that τ is proportional to L^2 at small L . We found that a twofold increase in L (from $L = 0.5$ cm to 1 cm) indeed resulted in about a fourfold increase in τ (see Fig. 4 legend), but further increase in L (from 1 cm to 2 cm) increased τ by about a factor of seven. This remains to be explained theoretically. □

Microbially mediated cerium oxidation in sea water

James W. Moffett

Chemistry Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

REDOX processes influence the geochemistry of many elements in the ocean, but attributing the distribution of these elements in the water column and sediments to specific redox processes is difficult because they are also influenced by non-redox processes and by other inputs that are poorly constrained. Cerium provides an opportunity to study redox processes in the ocean¹⁻⁵ because its redox chemistry leads to its enrichment or depletion ('cerium anomalies') with respect to its lanthanide neighbours. A detailed understanding of Ce geochemistry is lacking, however, because of the paucity of knowledge of the redox rates and mechanisms in natural waters. Here I report measurements of Ce(III) oxidation rates using radiotracers in seawater samples collected in the Sargasso Sea and in Vineyard Sound, Massachusetts, which are much faster than previous, indirect estimates¹. The data indicate that the negative Ce anomaly in sea water is the result of microbial oxidation followed by preferential scavenging of Ce(IV); no abiotic oxidation was detectable. This suggests that inhibition by sunlight of microbial oxidation and scavenging of Ce and Mn contribute to the pronounced surface maxima observed for these elements.

Current knowledge of cerium redox chemistry in natural waters has been inferred from its distribution in the water column and various mineral phases, and from thermodynamic data in the literature¹⁻⁵. These indicate that Ce is preferentially removed from oxic waters by oxidation to insoluble Ce(IV), the thermodynamically stable form. Manganese oxide phases, such as Mn nodules, are strongly enriched in Ce (refs 2, 6), indicating that oxidation on Mn oxides may be a mechanism for Ce removal. The similarity of Ce and Mn marine geochemistry is also consistent with this hypothesis. But it is not clear if this similarity arises because of Ce oxidation on Mn oxides or because the distributions of both elements are controlled by common processes. Existing data provide no information on the timescales of the redox transformations of Ce, essential for complete understanding of its geochemistry and usefulness as a redox tracer.

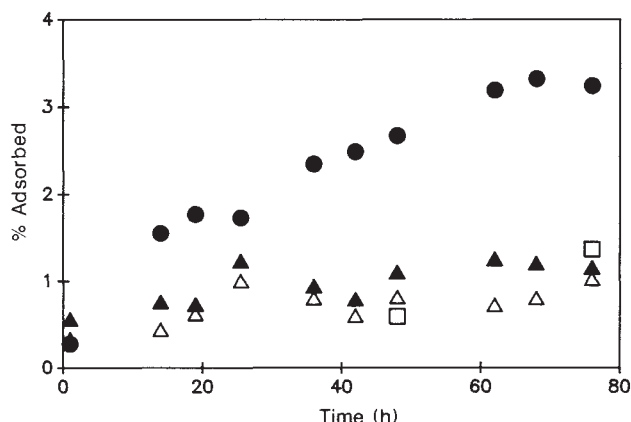


FIG. 1 Fraction of added radiotracer adsorbed on 0.2- μ m-filterable particles for a sample collected at 120 m in the Sargasso Sea (32°16'N, 64°36'W). Samples were collected in Teflon-coated bottles on Kevlar wire and processed within 2 h of collection. All samples were incubated in the dark, at 25 °C in 2-l polycarbonate bottles. Adsorption of tracers on container walls was negligible. Filters counted in NaI detector. Absolute precision = $\pm 0.2\%$. ●, ¹³⁹Ce; ▲, ¹³⁹Ce + 1 g l⁻¹ azide (added to entire sample 1 h before tracer addition); □, ¹³⁹Ce + 10⁻⁴ M ascorbate (ascorbate solution, pH 8, added to each aliquot 1 h before filtration); △, ¹⁵²Eu.

Received 15 September 1989; accepted 20 March 1990.

1. Winfree, A. T. & Strogatz, S. H. *Nature* **311**, 611-615 (1984).
2. Winfree, A. T. *Science* **181**, 937-939 (1973).
3. Welsh, B., Gomati, S. & Burgess, A. *Nature* **304**, 611-614 (1983).
4. Jahnke, W., Henze, C. & Winfree, A. *Nature* **336**, 662-665 (1988).
5. Medvinsky, A. B., Panfilov, A. V. & Pertsov, A. M. in *Self-Organization. Autowaves and Structures Far from Equilibrium* (ed. Krinsky, V.) 195-199 (Springer, Berlin, 1984).
6. Mikhailov, A. S., Panfilov, A. V. & Rudenko, A. N. *Phys. Lett. A* **109**, 246-250 (1985).
7. Panfilov, A. V., Rudenko, A. N. & Pertsov, A. M. *Dokl. Akad. Nauk SSSR* **279**, 1000-1002 (1984).
8. Biktashev, V. N. *Physica D* **36**, 167-172 (1989).
9. Keener, J. P. *Physica D* **31**, 269-276 (1988).
10. Braznik, P. K., Davydov, V. A., Zykov, V. S. & Mikhailov, A. S. *Abstr. 2nd All-Union Conference on Mathematical and Computational Methods* 118-119 (ONTI NCBI, Pushchino, 1987).
11. Zhabotinsky, A. M. & Zaikin, A. N. in *Oscillatory Processes in Biological and Chemical Systems* Vol. 2, 279-283 (ONTI NCBI, Pushchino, 1971).
12. Agladze, K. I. & Krinsky, V. I. *Nature* **296**, 424-426 (1982).
13. Allesio, M. A., Bonke, F. I. M. & Schopman, F. J. G. *Circul. Res.* **33**, 54-62 (1973).
14. Bures, J. & Gorelova, N. A. *J. Neurobiol.* **14**, 353-363 (1983).
15. Gerish, G. *Curr. Topics dev. Biol.* **3**, 157 (1968).
16. Panfilov, A. V. & Pertsov, A. M. *Dokl. Akad. Nauk SSSR* **274**, 1000-1002 (1984).
17. Kuhnert, L. *Naturwissenschaften* **70**, 464-466 (1983).
18. Nosztzious, Z., Hopfner, W., McCormick, W. D., Swinney, H. L. & Tam, W. Y. *Nature* **329**, 619-620 (1987).
19. Agladze, K. I., Krinsky, V. I. & Pertsov, A. M. *Nature* **308**, 834-835 (1984).

ACKNOWLEDGEMENTS. We thank K. I. Agladze and V. N. Biktashev for helpful discussions.

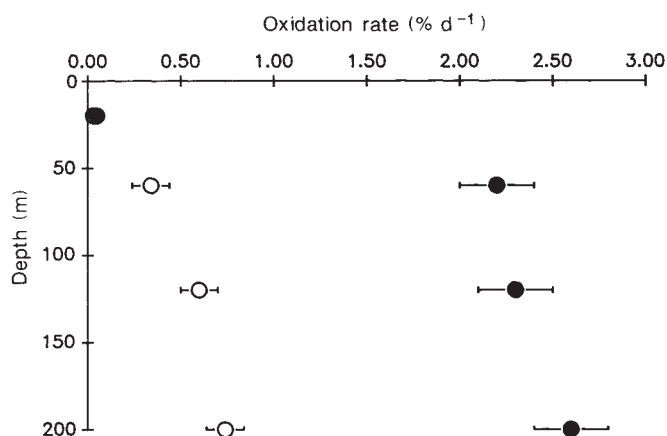


FIG. 2 Specific oxidation rates (% d⁻¹) of Ce(III) (○) and Mn(II) (●) as a function of sample depth. Calculated by subtracting ascorbate-treated samples from non-treated samples.

I have measured Ce oxidation in two diverse marine locations. Oxidation rates were measured by adding ¹³⁹Ce at tracer levels (that is, [¹³⁹Ce] < [total Ce]) to freshly collected sea water, and by following its uptake onto filterable particles. At specified times, aliquots of solution were filtered through 0.2-μm filters and the γ-activity counted. ¹⁵²Eu was also added as a non-redox-reactive lanthanide control to distinguish Ce oxidation from Ce(III) adsorption. Eu(III) is, however, in the middle of the lanthanide series and may be less particle reactive than Ce(III) in sea water because of stronger carbonate complexation^{1,7}, so an independent method was required to distinguish between Ce(III) and Ce(IV). I established that ascorbate will reduce Ce(IV) to Ce(III) in sea water at ambient pH (unpublished data). Addition of 10⁻⁴M ascorbate to these samples resulted in a decrease of particulate ¹³⁹Ce through reduction and solubilization of Ce(IV). Because ascorbate is not a chelator and the pH is unchanged, Ce(III) is unaffected. Therefore, Ce(IV) was determined as the difference in particulate ¹³⁹Ce in the presence and absence of ascorbate. ¹⁵²Eu was not affected by ascorbate addition in Sargasso Sea samples. Ascorbate has been used in the same way to distinguish Mn oxides from particulate Mn(II) (ref. 8). Figure 1 shows results for a sample collected from 120 m at the Sargasso Sea station. Uptake of ¹³⁹Ce and ¹⁵²Eu was rapid in the first hour. Thereafter, Eu uptake was slow, approaching a steady-state value for the rest of the experiment. The uptake of ¹³⁹Ce continued to increase roughly linearly throughout the experiment, although at a slower rate than the initial uptake. When ascorbate was added to some aliquots and allowed to react for one hour before filtration in these aliquots, most of the particulate ¹³⁹Ce was re-dissolved and the remaining particulate fraction was comparable to ¹⁵²Eu. Therefore Ce(III) and Eu(III) uptake were comparable, and the preferential Ce uptake was the result of Ce(III) oxidation.

The oxidation of Mn(II) is microbially mediated in sea water, and the process is inhibited in the presence of azide^{8,9}. Sodium azide (1 g l⁻¹) was added to samples from 120 m and incubations performed in parallel. With this treatment, uptake was comparable to that for ¹⁵²Eu and ¹³⁹Ce in the presence of ascorbate, indicating that no oxidation occurred (Fig. 1). Therefore the Ce(III) oxidation was microbial and not due to abiotic oxidation. Specific oxidation rates of Ce(III) (in % per day) as a function of depth are shown in Fig. 2, along with Mn(II) oxidation rates, which were measured using ⁵⁴Mn. Rates were very slow at 20 m but increased dramatically below the mixed-layer depth (25 m). The trend with depth was the same for Ce and Mn, with rates 3–4 times greater for Mn. The Mn results are in good agreement with an earlier study in the southwestern Sargasso⁸.

Uptake of ¹³⁹Ce and ¹⁵²Eu in Vineyard Sound is shown in Fig. 3. Only about half of the particulate Ce that can be re-dissolved by ascorbate was inhibited by azide addition (Fig. 3a), which might indicate that the remaining fraction was oxidized abiotically. But the Eu data (Fig. 3b) suggests an alternative explanation. Uptake of Ce and Eu is similar for both elements and the fraction of adsorbed ¹⁵²Eu redissolved after ascorbate addition was equivalent to the fraction of ¹³⁹Ce redissolved that is not affected by azide. Therefore this fraction was probably Ce(III) and Eu(III) adsorbed on Mn oxides, which were dissolved by ascorbate, rather than Ce(IV). The concentration of particulate Mn at the beginning of the experiments was 18 nM in Vineyard Sound (compared with <1 nM in Sargasso samples), which was evidently sufficient for trivalent lanthanide adsorption to be significant. These data provide evidence that abiotic oxidation on Mn oxides is unimportant relative to microbial oxidation, even when they adsorb a significant fraction of Ce(III).

The Mn oxidation rate in Vineyard Sound was 3.9% h⁻¹. The microbial Ce oxidation rate (obtained by subtracting the uptake curve for the azide-containing solution from the azide-free solution in Fig. 3a and measuring the initial slope) was 1.3% h⁻¹. Therefore, microbial Mn oxidation was three times faster than the microbial Ce oxidation, similar to the Sargasso Sea results. That oxidation rates for Ce(III) and Mn(II) are within a factor of three is remarkable given the completely different speciation of each element in sea water—Mn(II) exists predominantly as Mn_{aq}²⁺ and Ce(III) as strong carbonate complexes^{7,10}. Furthermore, comparable ratios of Mn to Ce oxidation were found in both locations, spanning a 100-fold range in absolute rates. This indicates that the same microbial mechanism is involved throughout the environments studied. The close and constant relationship between oxidative removal pathways for Ce and Mn in diverse marine environments such as those studied here undoubtedly contributes to the similarities in the elements' geochemistries. No data were obtained from deep or bottom waters, but microbial Mn oxidation occurs throughout the water column, so it is likely that Ce oxidation also does.

Many trace elements are highly depleted in surface waters, but depth profiles of Ce and Mn frequently display sharp surface concentration maxima, with levels declining rapidly below the mixed layer^{1,11–15}. For Mn this has been attributed to aeolian deposition^{11,12} and/or photoreductive dissolution of Mn oxides in surface waters and photo-inhibition of microbial Mn oxidation (leading to the accumulation of Mn(II) in the upper photic zone⁸). The relative importance of atmospheric deposition to the *in situ* redox processes is not clear. Some elements that have no redox chemistry (such as aluminium) can also display surface maxima¹⁶. Under certain well stratified conditions the residence time of thorium in the upper photic zone (0–50 m) is much greater than in the lower 50–100 m. As Th has no redox chemistry, this has been attributed to differences in particle flux resulting from differences in the plankton community¹⁷. Therefore, the Mn maxima could arise from these processes and not redox reactions. But Ce frequently displays unique maxima not observed for the trivalent lanthanides, and it is not preferentially enriched in aeolian dust^{18,19} so aeolian deposition cannot account for the maxima. Similarly, non-redox processes shown to be important for Th can also be ruled out because they would presumably lead to maxima for the other lanthanides. The unique maxima for Ce must involve its redox chemistry. At the Sargasso Sea station, low Ce oxidation rates measured in the upper photic zone would lead to a much longer residence time than in the lower photic zone, where oxidation rates are 30 times greater. This difference could lead to surface maxima. These data do not explain why oxidation at 20 m is so much slower than it is lower in the photic zone, but earlier work⁸ indicates that microbial Mn oxidation is photo-inhibited, which is probably the case for Ce also. Photoreduction of Ce(IV) may also be a factor, but this has not been investigated in this study.

Given the apparent similarities between Ce and Mn, it can be argued that *in situ* redox chemistry must be of prime importance in contributing to surface Mn maxima.

Cerium oxidation is faster in Vineyard Sound than in the Sargasso Sea, but it is relatively less important than Ce(III) uptake. After 40 h, 70% of the ^{139}Ce uptake is the result of Ce oxidation in the Sargasso Sea at 120 m but only 30% in Vineyard Sound. Furthermore, because Ce(III) uptake occurs on a shorter timescale than Ce(III) oxidation, oxidation is relatively less important in environments where particle residence times are short. Therefore, Ce anomalies should be less pronounced in areas of high particle flux than low flux, and this is generally the case. Estuarine and coastal regions have little or no Ce anomaly, whereas oligotrophic waters have large negative Ce anomalies^{1,4,14,15}. Adsorption kinetic data obtained from radio-tracer experiments may be incorporated into non-equilibrium models of trace-element scavenging. For many elements (such as thorium, tin, scandium and zinc) this is difficult because of complex adsorption kinetics²⁰. As the Ce anomaly seems to arise from a unique process with relatively straightforward kinetics, development of a Ce scavenging model may be a useful starting point for extension to more-complicated systems. Such a model would also allow more accurate predictions of the response of the Ce anomaly to changes in oceanic redox conditions. □

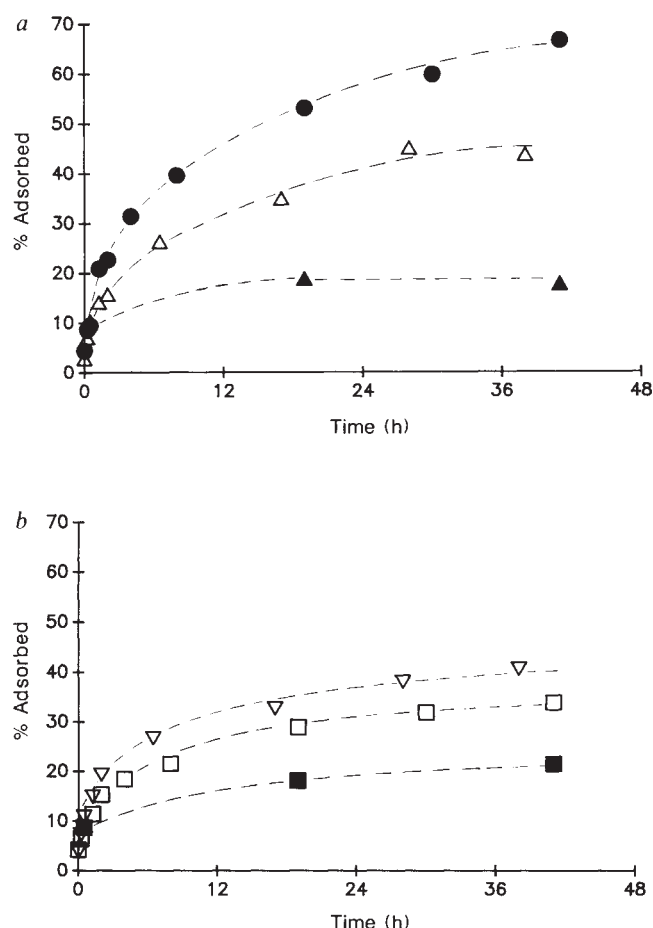


FIG. 3 Fraction of added radiotracers adsorbed on 0.2- μm -filterable particles in Vineyard Sound samples. Samples collected from a small boat in a Pyrex carboy at a location 7 km southeast of Woods Hole (10 August 1988). Samples were treated as for Sargasso samples. a, ●, ^{139}Ce ; △, ^{139}Ce + azide; ▲, ^{139}Ce + ascorbate. b, □, ^{152}Eu ; ▽, ^{152}Eu + azide; ■, ^{152}Eu + ascorbate. Absolute precision = $\pm 5\%$. Eu uptake was significantly higher in azide-treated samples than in non-treated samples. This may be the result of alteration of particulate matter by azide or may reflect real differences between samples. Differences are small relative to other differences in this figure.

Received 21 December 1989; accepted 9 April 1990.

1. Elderfield, H. *Phil. Trans. R. Soc. Lond. A* **325**, 105–126 (1988).
2. Goldberg, E. D., Koide, M., Schmitt, R. A. & Smith, R. H. *J. geophys. Res.* **68**, 4209–4217 (1963).
3. Elderfield, H. & Pagett, R. *Sci. Tot. Environ.* **49**, 175–197 (1986).
4. Sholkovitz, E. R. & Elderfield, H. *Global Biogeochem. Cycles* **2**, 157–176 (1988).
5. Murphy, K. & Dymond, J. *Nature* **307**, 444–447 (1984).
6. Elderfield, H., Hawkesworth, C. J., Greaves, M. J. & Calvert, S. E. *Geochim. cosmochim. Acta* **45**, 1231–1234 (1984).
7. Cantrell, K. J. & Byrne, R. H. *Geochim. cosmochim. Acta* **51**, 597–605 (1987).
8. Sunda, W. G. & Huntsman, S. A. *Deep Sea Res.* **35**, 1297–1317 (1988).
9. Tebo, B. M. & Emerson, S. *Biogeochem.* **2**, 149–161 (1982).
10. Turner, D. R., Whitfield, M. & Dickson, A. G. *Geochim. cosmochim. Acta* **45**, 855–882 (1981).
11. Landing, W. & Bruland, K. W. *Geochim. cosmochim. Acta* **51**, 29–43 (1987).
12. Martin, J. H. & Knauer, G. A. *Earth planet Sci. Lett.* **51**, 266–274 (1980).
13. De Baar, H. J. W., Bacon, M. P. & Brewer, P. G. *Nature* **301**, 203–204 (1983).
14. De Baar, H. J. W., Bacon, M. P., Brewer, P. G. & Bruland, K. W. *Geochim. cosmochim. Acta* **49**, 1949–1953 (1985).
15. German, C. thesis, Cambridge Univ. (1988).
16. Orians, K. J. & Bruland, K. W. *Nature* **316**, 427–429 (1984).
17. Coale, K. H. & Bruland, K. W. *Limnol. Oceanogr.* **32**, 189–200 (1987).
18. Rahn, K. A. *The Chemical Composition of Atmospheric Aerosol*, Tech. Rep. 265 (Univ. of Rhode Island, 1976).
19. Schneider, B. *J. geophys. Res.* **90**, 10744–10746 (1985).
20. Jannasch, H. W., Honeyman, B. D., Balistrieri, L. S. & Murray, J. W. *Geochim. cosmochim. Acta* **52**, 567–577 (1988).

ACKNOWLEDGEMENTS. I thank E. Sholkovitz and W. Sunda for discussions. This work was supported by the NSF.

Ocean-ridge basalt compositions correlated with palaeobathymetry

M. J. Keen*, E. M. Klein† & W. G. Nelson‡

* Atlantic Geoscience Centre, Geological Survey of Canada, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada, B2Y 4A2

† Department of Geology, Duke University, Durham, North Carolina 27708, USA

‡ Smithsonian Institution, Washington, DC 20560, USA

KLEIN and Langmuir¹ investigated the relationships between the chemistry and water depth of eruption of zero-age mid-ocean-ridge basalts around the world. They showed that the regionally averaged values of the major-element oxides Na_2O and FeO , corrected for low-pressure fractionation, and the ratio $\text{CaO}/\text{Al}_2\text{O}_3$, correlate with their axial depths. Klein and Langmuir pointed out that the same principles should also apply to older ocean crust, where the relationships that they had established could be used as 'calibration curves', and that there was a "possibility of petrologically constraining the origins of bathymetric anomalies throughout the ocean basins"¹. Keen² showed that this can be achieved if the initial depths of eruption of older basalts are found by correcting their present depths below sea level for sediment loading and thermal subsidence. Here we show that the chemistry of basalts from older oceanic crust in the Atlantic and Indian Oceans correlates with their depth restored in this way to the original water depth at the time of eruption.

The Ocean Drilling Program's Hole 647A in the extinct mid-ocean ridge of the Labrador Sea illustrates our approach (Figs 1 and 2). Basalts were encountered at 4,561 m below sea level, beneath 699 m of sediment^{3,4}. They are typical mid-ocean-ridge basalts (MORB), 56 Myr old, but still fresh⁵. Values of Na_2O and FeO corrected for low-pressure fractionation to a common MgO value of 8 wt%, $\text{Na}_{8.0}$ and $\text{Fe}_{8.0}$ respectively, and the ratio $\text{CaO}/\text{Al}_2\text{O}_3$, have been calculated for each of the analyses⁵ that meets the criteria established by Klein and Langmuir¹ in terms of the range of acceptable values of MgO wt% (Table 1).

We found the initial depth of eruption of the basalts by 'back-tracking'⁶. Sediments have depressed the ocean crust at site 647A, and so the basalts have been 'unloaded' assuming point-wise Airy isostasy⁷, using a mean value of sediment density calculated from the set of measurements for the site³ (Table 1). The ocean crust at site 647A has also subsided thermally, and this subsidence has also been restored. We assume that the average depth at which ocean crust is formed is 2,500 m (ref. 8).

TABLE 1 Sites on mid-ocean ridges

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Site	Lat. (deg)	Long. (deg)	Depth (m)	Sediment (m)	Igneous section (m)	Basalt age (Ma)	Restored (m)	Unloaded (m)	Now (m)	Source glass rock	Na _{8.0}	Fe _{8.0}	Number of Na, Fe analyses	CaO/Al ₂ O ₃	Number of Ca/Al analyses	Refs
14	28.33 S	20.94 W	4,346	107	1.5	39	2,229	4,415	4,453	SG	—	—	—	0.740	1	20
15	30.89 S	17.98 W	3,938	141	1.4	21	2,465	4,031	4,079	SG	2.30	8.92	1	0.837	1	20
18	27.99 S	08.01 W	4,022	178	0.3	26	2,363	4,148	4,200	SG	—	—	—	0.777	1	20
19	28.54 S	23.68 W	4,685	141	4.3	51	2,284	4,783	4,826	SG	2.20	9.50	1	0.826	1	20
105	34.90 N	69.17 W	5,245	623	9	156	2,044	5,676	5,875	SG	2.06	10.45	1	0.804	1	21–24
335	37.30 N	35.20 W	3,188	454	108	13	2,244	3,506	3,642	SG	2.40	9.45	1	0.761	1	24–27
350	67.06 N	08.29 W	1,275	362	18	41	–668	1,566	1,637	DR	1.89±0.04	9.64±0.12	3	0.767±0.010	3	28
407	63.94 N	30.58 W	2,472	304	155	35.5	629	2,714	2,776	SG	1.78	10.25	1	0.817	1	24, 29
409	62.62 N	25.95 W	832	80	239	2.4	352	894	912	DG	1.63±0.20	9.17±0.09	2	0.835±0.023	2	24, 29
417D	25.11 N	68.04 W	5,482	343	366	118	2,320	5,733	5,830	SG	2.09±0.10	9.92±0.32	22	0.825±0.016	22	30, 36
418A	25.04 N	68.06 W	5,511	324	544	118	Combine 417 and 418									
558	37.77 N	37.34 W	3,754	408	154	37	1,873	4,002	4,162	SG	2.27±0.13	9.67±0.05	2	0.755±0.016	3	31
559	35.12 N	40.92 W	3,754	238	63	35	1,857	3,927	3,992	SG	2.43	9.62	1	0.725	1	31
563	33.64 N	43.77 W	3,786	365	19	35.5	1,948	4,033	4,151	SG	2.16±0.01	10.02±0.18	3	0.833±0.027	3	31
564	33.74 N	43.77 W	3,820	284	81	35.5	1,940	4,025	4,104	SG	2.17	10.06	1	0.773	1	31
647A	53.33 N	45.26 W	3,862	699	37	56	1,692	4,311	4,561	DR	1.93±0.18	11.14±0.69	9	0.828±0.033	14	3–5
765D	15.98 S	117.58 E	5,724	924	271	139	2,720	6,270	6,648	DR	2.05±0.23	9.24±1.10	10	0.805±0.030	10	32, 33
All	23.42 N	52.03 W	5,615	<100	D	47	3,216	5,615	5,615	SG	2.57	8.52	1	0.691±0.038	2	34, 35

(1) Data. Rationale: Mid-Atlantic Ridge and its satellites—a coherent body of data; older than 20 Myr so that subsidence would be significant; 335, from our initial pilot study; 765D (Indian Ocean), deep now; All, deep formerly; 350, 407 and 409, shallow now and formerly. The data shown here form the complete Smithsonian data base for Atlantic glasses older than 20 Myr, supplemented by whole-rock data and DSDP glass data for extremes in depth. No data analysed have been excluded. (6) Section drilled through igneous rocks. D, dredge haul. (7) Age: ref. 36 throughout. (8) and (9), Density: sea water, 1.03 g cm⁻³; mantle, 3.33 g cm⁻³. (11) Source: D, Deep Sea Drilling Project/Ocean Drilling Program volumes; S, Smithsonian Volcanic Glass Project Data; G, glass; R, whole rock. (12)–(16) Criteria for inclusion in Table 1 and in plots. All data: meet MgO criteria of ref. 1. Smithsonian Volcanic Glass Project Data: inclusion in that data base. Glass analyses, DSDP: 409: ref. 37, Table 3, samples 409 7 6b, 409 13 2; both glasses; K₂O mean 0.15. We observe that the average of the two totals is 97.80. Whole-rock analyses: DSDP/ODP: 350: ref. 28, p. 661, Table 3, three analyses: H₂O (total) average 1.16; K₂O average, 0.28; total average, 99.27. 647A: ref. 5, Table 1: analyses for Site 647A: total average, 99.38; K₂O average, 0.05; average H₂O + 1.38 (of eight analyses reporting). 765D: ref. 32, analyst T. Plank: Table 765-P-2, analyses with 0.0 < LOI < 0.6; these had: total average, 99.52; K₂O average, 0.50; LOI average, 0.33. LOI is loss on ignition. (12), (13) and (15), Fig. 3: correction for low-pressure fractionation: see refs 1, 2: Na_{8.0} = Na₂O + 0.373 × (MgO) – 2.98; Fe_{8.0} = FeO + 1.664 × (MgO) – 13.313. Na_{8.0} and Fe_{8.0} are calculated for samples with 5.0–8.5% MgO. CaO/Al₂O₃ are calculated for samples with >5.0% MgO. Fe_{8.0} = 2.033 × (5.233 – b_{Fe}), where b_{Fe} = Na_{8.0} – 0.19 × Fe_{8.0}. (14) and (16), Numbers of analyses: numbers of analyses available that passed screening.

After 56 Myr, crust formed at this water depth should have subsided to 5,119 m (Table 1). But the depth of the basalts at site 647A corrected for unloading (4,311 m) is 808 m less than this. Assuming that initial differences in elevation are preserved as the crust subsides overall, the initial depth of the crust at site 647A becomes 1,692 m, not 2,500 m. This value is consistent with the generally shallow crust in the Labrador Sea, attributed to the effects of a hotspot that developed about 60 Myr ago^{4,9–11}

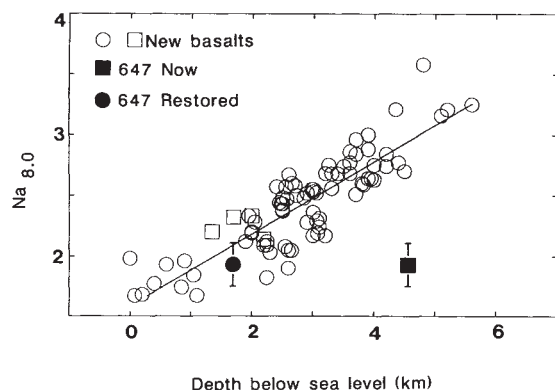


FIG. 1 The average Na_{8.0} value of the 56-Myr-old basalts from Ocean Drilling Program site 647A (Na₂O wt% corrected for low-pressure fractionation¹) plots in the field established by Klein and Langmuir¹ for zero-age axial basalts when restored to the initial water depth of site 647A (●), but not at the present depth below sea level of the basalt (■). ○, □ ('new'): ○, values for present-day ocean ridge basalts world-wide¹; □, values described by Klein and Langmuir¹ as 'anomalous' from various axial hotspot centres. ■ ('647 now'): average value of Na_{8.0} for fresh basalts from site 647A plotted at the present depth below sea level of the basalts. ● ('647 restored'): the same value plotted at the depth of the basalts restored to the initial depth at the time of formation by correcting for sediment loading and thermal subsidence (see text). Bars: one standard deviation about the mean Na_{8.0} value. Lines: linear regression on the open circles plotted solely as an aid to the eye.

beneath the Davis Strait to the north or beneath eastern Greenland.

At their present depth below sea level, the Na_{8.0} value for basalts from site 647A does not plot in the field defined by Klein and Langmuir¹. But the value does plot in this field after the site is restored to its initial water depth by correcting for the effects of loading and thermal subsidence (Fig. 1).

The chemistry of older basalts from the other sites that we studied in the Atlantic and Indian Oceans shows the same relationship to the initial depth as it does at site 647A (Table 1; Figs 2 and 3). Figure 3 shows that values of Na_{8.0}, CaO/Al₂O₃ and Fe_{8.0} in general fall within the fields defined by Klein and Langmuir¹ when the basalts are restored to their original water depths at the time of crustal formation. Furthermore, Fig. 3 shows that the relationship between chemistry and initial water depth at these sites holds over at least a substantial part of the range delineated by Klein and Langmuir¹, including shallow sites close to known hotspots such as Iceland¹.

Although the data for older basalts restored to initial water depths in general fall within the fields defined by Klein and Langmuir¹, there is only a weak inverse correlation between restored depth and Fe_{8.0}. This is consistent with the weaker correlation for zero-age basalts between regional averages of depth and Fe_{8.0} compared with the correlations between depth and Na_{8.0} or CaO/Al₂O₃ (refs 12 and 13). Klein and Langmuir¹² suggested that because iron abundances primarily reflect the pressure of melting, the weaker correlation for Fe_{8.0} may reflect in part the sampling of instantaneous melts from different pressures within a zone of melting. They proposed an algorithm to isolate the component of iron variation that reflects regional differences in solidus pressure and noted a good inverse correlation between this parameter (Fe_{8.0}^G) and depth for individual analyses (Fig. 3g, h). The data for individual analyses for older basalts also show a good correlation between Fe_{8.0}^G and restored depth, and fall within the field of data presented by Klein and Langmuir¹².

In an approach such as this, uncertainty may arise from several sources. First, the chemical composition of ocean basalts is often altered by interactions with sea water. We therefore used analy-

ses of fresh glasses where we could, from the Smithsonian Volcanic Glass Data Base¹⁴ and the Deep Sea Drilling Project (see refs in Table 1). These have been supplemented by whole-rock analyses of fresh rocks for the three sites representing extremes in depth where glass analyses were not available (Table 1); these bulk analyses, however, may not reflect liquid compositions because of the presence of phenocrysts in some samples. We have attempted to estimate chemical variability among samples from a single site using the standard deviation about the average values at sites where more than one group of samples was available (Fig. 3, Table 1).

Uncertainty also arises in restoration to initial depth because of the corrections for sediment loading, the effects of errors in crustal age on estimates of thermal subsidence, and the assumption that initial differences in elevation at ridge crests are preserved as the crust subsides. The correction for sediment loading depends on the average density of the sediment column⁶, and this can be estimated from bulk measurements of density, the sediment thickness⁷, or the two-way reflection time at the site¹⁵. Comparisons of the three methods at the sites that we used show that this uncertainty is in the range of only tens of metres. Similarly, errors in crustal age are usually sufficiently small that we can neglect the uncertainty they cause in restoring thermal subsidence.

The assumption that initial elevations are preserved is well established for 'normal' mid-ocean ridges and for aseismic ridges⁶. It provides a clear reference for the sites used here, and independent evidence supports many of the initial depths that we have predicted. For example, the crest of the Mid-Atlantic Ridge along the flow line east of the site AII is deeper than normal at present¹⁶, and other studies indicate that sites near Iceland (such as 350, 407, 409) and sites younger than about 36 Myr near the Azores (559, 563, 564) should be shallow, as we find¹⁷. Probing the details of such depth anomalies will be interesting; for example, the original depths of sites now on swells will depend on the timing of their formation with respect to that of the swell itself.

Consequently, the correlations we see between restored initial eruption depths and major-element geochemistry—the right-

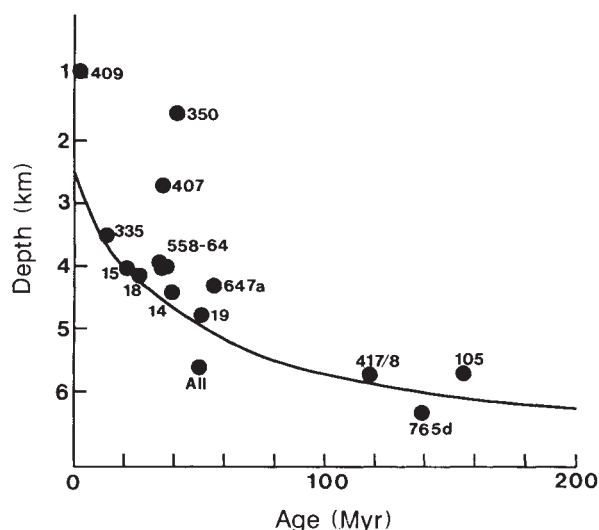


FIG. 2 The sites used in this study plotted at their unloaded depths and ages. The unloaded depth is the present depth of the basalt at a site below sea level, corrected for loading by sediments. The solid curve is the average subsidence curve for the ocean crust⁸. All locations represent Deep Sea Drilling Project or Ocean Drilling Program site numbers, except for AII—a dredge haul from the western North Atlantic³⁴. See Table 1 for locations.

hand panels of Fig. 3—lead us to suggest that our observations on basalts from older oceanic crust in the Atlantic Ocean (and from the one site in the Indian Ocean) support and extend the hypothesis of Klein and Langmuir¹ over a wide range of initial depths and age. The hypothesis seems to have survived a rather robust test.

Klein and Langmuir¹ noted a correlation between crustal thickness and basalt chemistry, and suggested that the covariations among zero-age depth, chemistry and crustal thickness result from variations in subsolidus mantle temperatures. Efforts to relate oceanic crustal thickness to other physical parameters have only rarely been successful^{18,19}. Our results indicate that

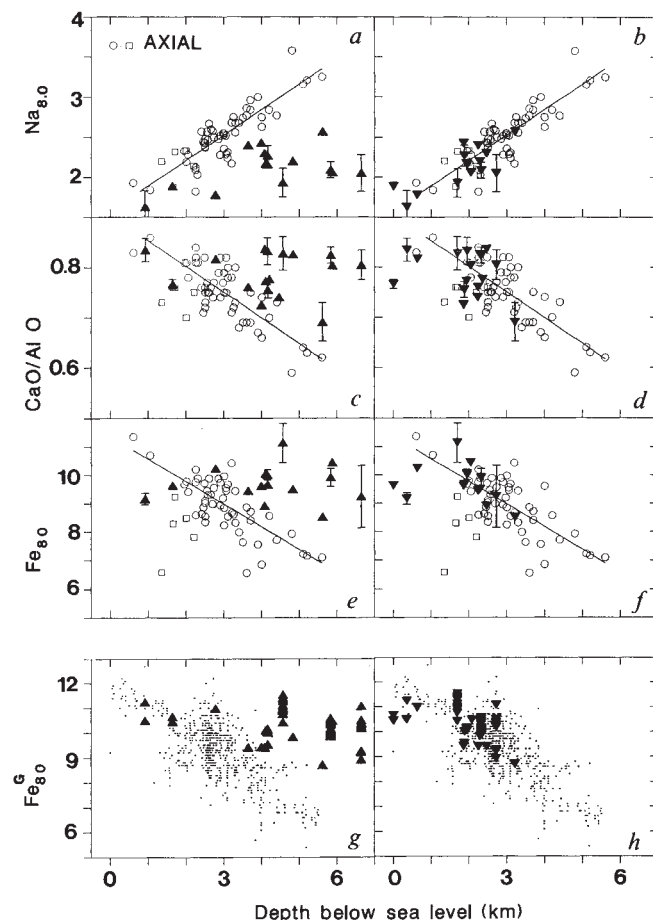


FIG. 3. The major-element geochemistry of older basalts from the mid-ocean ridges of the ocean basins correlates with the original water depth of the basalts (the depth when the basalts erupted). The variations in averaged $\text{Na}_{8.0}$ and $\text{Fe}_{8.0}$, in averaged $\text{CaO}/\text{Al}_2\text{O}_3$, and in individual analyses of $\text{Fe}_{8.0}^G$ fall within the fields defined by Klein and Langmuir^{1,12} for zero-age axial basalts after restoration to their original depths. $\text{Fe}_{8.0}^G$ is the iron variation calculated to maximize the component of iron that reflects solidus pressure¹² (see text). a, c, e, g, $\text{Na}_{8.0}$, $\text{CaO}/\text{Al}_2\text{O}_3$, $\text{Fe}_{8.0}$, $\text{Fe}_{8.0}^G$, respectively. Values from older basalts from the oceanic ridges of the Atlantic Ocean and Indian Ocean (one site) (Table 1) with standard deviations where available, plotted at their present depths below sea level ($\blacktriangle$, 'present depth'). They do not fall into the fields defined by the modern axial basalts ($\circ$, $\square$, dots). b, d, f, h, $\text{Na}_{8.0}$, $\text{CaO}/\text{Al}_2\text{O}_3$, $\text{Fe}_{8.0}$, $\text{Fe}_{8.0}^G$, respectively. Values from older basalts from the oceanic ridges restored to their original water depths at the time of formation by correcting present basement depths for loading by sediments and for thermal subsidence (∇ , 'restored depths'). They fall within the same fields as the zero-age axial basalts. Lines: linear regression on the open circles, plotted only as an aid to the eye comparing left and right panels. a–f, $\circ$ and $\square$, ('axial'); regional averages of zero-age axial basalts^{1,12} (see caption to Fig. 1). g and h, dots: individual analyses of $\text{Fe}_{8.0}^G$ for zero-age axial basalts¹².

we should extend the approach that we have devised here, and explore the relationships between crustal thickness and restored initial depth. Crustal thickness may perhaps be established more reliably on older crust than on younger crust, and so we may have opened up the possibility of more fruitful investigations of the relationships between depth, crustal thickness and crustal composition than have been possible before. Klein and

Langmuir¹ also suggested that temporal variations in mantle temperatures may be explored by sampling older oceanic crust along a flow line emanating from a particular ridge segment. Our results show that the correlation between zero-age chemistry and depth can indeed be extended to older oceanic basalts, and make it attractive to seek systematic temporal variations in mantle temperature from studies of older crust. □

Received 8 January; accepted 20 March 1990.

1. Klein, E. M. & Langmuir, C. H. *J. geophys. Res.* **92**, 8089–8115 (1987).
2. Keen, M. J. *Trans. Am. geophys. Un.* **70**, 1400 (1989).
3. Srivastava, S. P. *et al. Proc. Ocean Drilling Program, Init. Rep.* **105**, 675–725 (1987).
4. Srivastava, S. P. & Arthur, M. A. *Proc. Ocean Drilling Program, Scientific Results* **105**, 989–1089 (1989).
5. Clarke, D. B., Cameron, B. I., Muecke, G. K. & Bates, J. L. *Can. J. Earth Sci.* **26**, 956–968 (1989).
6. Sclater, J. G., Meinke, L., Bennett, A. & Murphy, C. *Mem. geol. Soc. Am.* **163**, 1–19 (1985).
7. Le Douarin, S. & Parsons, B. *J. geophys. Res.* **87**, 4715–4722 (1982).
8. Parsons, B. & Sclater, J. G. *J. geophys. Res.* **82**, 803–827 (1977).
9. Hyndman, R. D. *Can. J. Earth Sci.* **10**, 637–644 (1973).
10. Keen, C. E. & Barrett, D. L. *Geophys. J. R. astr. Soc.* **30**, 253–271 (1972).
11. White, R. & McKenzie, D. *J. geophys. Res.* **94**, 7685–7730 (1989).
12. Klein, E. M. & Langmuir, C. H. *J. geophys. Res.* **94**, 4241–4252 (1989).
13. Brodholt, J. P. & Batiza, R. *J. geophys. Res.* **94**, 4231–4239 (1989).
14. Nelson, W. G. & O'Hearn, T. in *The Geology of North America, Volume M, The Western North Atlantic Region* (eds Vogt, P. R. & Tucholke, B. E.) 117–136 (Geol. Soc. Am., Boulder, 1986).
15. Crough, S. T. *J. geophys. Res.* **88**, 6449–6454 (1983).
16. Sclater, J. G., Lawver, L. A. & Parsons, B. *J. geophys. Res.* **80**, 1031–1052 (1975).
17. Sclater, J. G., Hellinger, S. & Tapscott, C. *J. Geol.* **85**, 509–552 (1977).
18. Purdy, G. M. & Ewing, J. in *The Geology of North America, Volume M, The Western North Atlantic Region* (eds Vogt, P. R. & Tucholke, B. E.) 313–330 (Geol. Soc. Am., Boulder, 1986).
19. Reid, I. & Jackson, H. R. *Mar. geophys. Res.* **5**, 165–172 (1981).
20. Maxwell, A. E. *et al. Init. Rep. DSDP Leg 3* (1970).
21. Hollister, C. D. *et al. Init. Rep. DSDP Leg 11* (1972).

22. Vogt, P. R. & Tucholke, B. E. (eds) *The Geology of North America, Volume M, The Western North Atlantic Region* (Geol. Soc. Am., Boulder, 1986).
23. Sclater, J. G. & Wixon, L. in *The Geology of North America, Volume M, The Western North Atlantic Region* (eds Vogt, P. R. & Tucholke, B. E.) 257–270 (Geol. Soc. Am., Boulder, 1986).
24. Bryan, W. B. & Frey, F. A. in *The Geology of North America, Volume M, The Western North Atlantic Region* (eds Vogt, P. R. & Tucholke, B. E.) 271–296 (Geol. Soc. Am., Boulder, 1986).
25. Aumento, F. *et al. Init. Rep. DSDP Leg 37* (1977).
26. Howe, R. C. *Init. Rep. DSDP Leg 37*, 909–916 (1977).
27. Miles, G. A. *Init. Rep. DSDP Leg 37*, 929–962 (1977).
28. Talwani, M., Udintsev, G. *et al. Init. Rep. DSDP Leg 38* (1976).
29. Luyendyk, B. P. *et al. Init. Rep. DSDP Leg 49* (1978).
30. Donnelly, T. *Init. Rep. DSDP Legs 51–53* (1979).
31. Bougault, H. *et al. Init. Rep. DSDP Leg 82* (1985).
32. Ludden, J. *et al. Proc. Ocean Drilling Program Init. Rep.* **123**, 230–289 (1990).
33. *JOIDES J.* **15**(1), 3–7 (1989).
34. Ludden, J. N. & Thompson, G. *Earth planet. Sci. Lett.* **43**, 85–92 (1979).
35. Klitgord, K. & Schouten, H. in *The Geology of North America, Volume M, The Western North Atlantic Region* (eds Vogt, P. R. & Tucholke, B. E.) 351–378 (Geol. Soc. Am., Boulder, 1986).
36. Kent, D. V. & Gradstein, F. M. in *The Geology of North America, Volume M, The Western North Atlantic Region* (eds Vogt, P. R. & Tucholke, B. E.) 45–50 (Geol. Soc. Am., Boulder, 1986).
37. Wood, D. A. *et al. Init. Rep. DSDP Leg 49*, 597–655 (1978).

ACKNOWLEDGEMENTS. We thank C. E. Keen, D. J. W. Piper, D. B. Clarke, F. Gradstein, R. Courtney, R. D. Hyndman, T. Plank, C. H. Langmuir, J. Cann, S. Srivastava, R. White, G. Thompson and J. Bogen for suggestions, staff of the Ocean Drilling Program Office and T. O'Hearn for assistance with electron microprobe analyses of glasses.

Seismic quiescence at Parkfield: an independent indication of an imminent earthquake

M. Wyss*, P. Bodin* & R. E. Habermann†

* Cooperative Institute for Research in Environmental Science/Department of Geology, University of Colorado, Boulder, Colorado 80309, USA

† National Geophysical Data Center of NOAA, Boulder, Colorado 80303, USA

THE Parkfield segment of the San Andreas fault midway between San Francisco and Los Angeles has experienced six moderate-sized earthquakes since 1850, with an average recurrence time of about 22 years. Based on this regular behaviour, and the occurrence of the most recent earthquake in 1966, a shock of local magnitude $M_L = 5.7$ is expected to occur by 1993^{1,2}. Here we suggest that a current lack of small earthquakes on this segment could be a precursory anomaly to the next characteristic Parkfield earthquake. Only four earthquakes of $M \geq 2.5$ have occurred on the segment since January 1986, whereas twenty would have been expected, extrapolating from the mean background rate. Comparison with other cases of precursory quiescence suggests that the $M = 5.7$ earthquake should occur in the next two years, rupturing the same 35-km fault segment that ruptured in 1966³.

Seismic quiescence is a phenomenon that has been used successfully to predict moderate and large earthquakes^{4–6}. Seismicity consists of two parts: events that depend on each other (clusters) and independent events. After removing the clusters from the data⁷, the remaining seismicity (background) shows a remarkably constant rate over many years. Against this constant rate a decrease by about a factor of two (seismic quiescence) that lasts for approximately a year or more is an unusual occurrence and is often followed by a mainshock⁸. The quiescence is thought to signal a change in the physical process governing the production of small earthquakes in the rupture volume.

Usually the quiescence outlines the mainshock rupture volume and affects all magnitude bands^{4–6,8–11} but it has been proposed that in some cases the rate of small-magnitude earthquakes remains constant^{12,13}.

An example of precursory seismic quiescence is shown in Fig. 1a. The Stone Canyon case was used as a basis for estimating what to expect in the Parkfield area because it occurred on the San Andreas fault about 80 km northwest of Parkfield and comes from a similar tectonic environment.

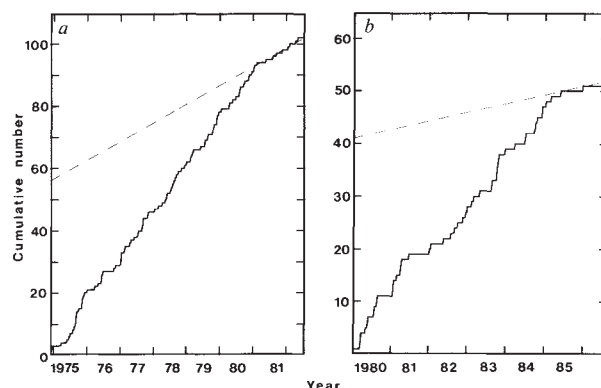


FIG. 1 Applicability of the quiescence hypothesis to the Parkfield segment of the San Andreas fault, tested by comparing the patterns of seismicity rate in the source volumes of mainshocks at a, Stone Canyon and b, Parkfield. The mainshocks ($M = 5$ at Stone Canyon, $M = 3.6$ at Parkfield) occurred at the time of the right end of the plots of cumulative seismicity (all shocks $M > 1.8$ at Stone Canyon and $M > 0.9$ at Parkfield) and quiescence started 1.5 years before the mainshocks. The slope of these curves is the mean seismicity rate. The dashed lines extrapolate the anomalously low precursory rate back in time for comparison with the observed background rate. The similarity of the pattern strongly suggests that the quiescence hypothesis should be as applicable at Parkfield as at Stone Canyon, where a successful prediction was made⁴.

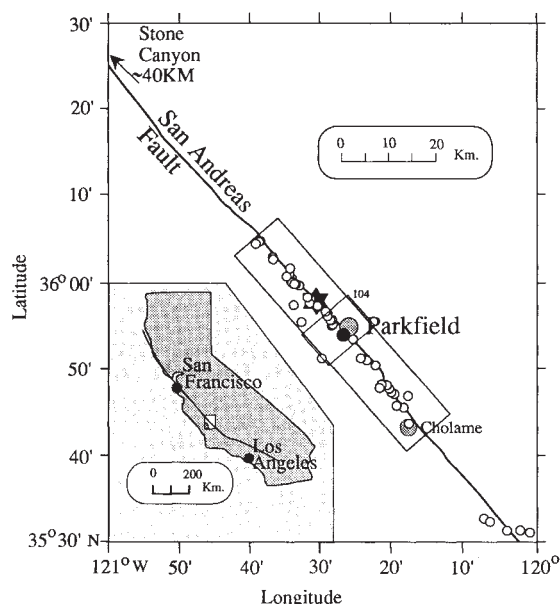


FIG. 2 Map of the Parkfield segment of the San Andreas fault showing the epicentres of earthquakes with $M > 2.4$ that occurred between 1975 and 1985. Since the beginning of 1986 only four such earthquakes occurred. The star and solid dot denote the mainshock epicentres of 28 June 1966 and 29 August 1986 respectively. The large rectangle outlines the area for which the cumulative seismicity curves are shown in Fig. 3, the small rectangle (104) outlines the area used for the earthquake count in Fig. 1b. The quiescence extends along the fault from latitude 35.7°N to 36.05°N approximately but is most pronounced north of Parkfield.

The macroseismic observations and surface-wave records seem to indicate that the Parkfield fault segment ruptures repeatedly in earthquakes with practically the same location, fault length and fault slip (a characteristic earthquake)¹. Based on the periodicity with which the last five such earthquakes have occurred¹, the next characteristic Parkfield earthquake was expected to occur between 1983 and 1993. This event has not taken place yet.

Seismicity near Parkfield is monitored by the US Geological Survey, which publishes an earthquake catalogue for the area outlined in Fig. 2. Artificially introduced rate changes through inadvertent changes of the magnitude scale or of the detection capability, which are present in all the catalogues we have examined, can sometimes be identified using the magnitude signature method^{9,10,14-18}. The details of our estimate of the minimum magnitude of homogeneous recording as a function of time, and the artificial magnitude shifts in the Parkfield catalogue will be given elsewhere.

The cumulative number of earthquakes for a subsegment of the fault is shown in Fig. 1b. The segment is 10 km long and randomly placed, covering part of the rupture of one of the two largest earthquakes that took place in the past ten years in the Parkfield area (29 August 1986, $M = 3.6$). The figure shows that quiescence preceded this event, and a more detailed analysis shows that it could have been predicted generating only one false alarm, using an alarm level from the Stone Canyon case history (Fig. 1a). We therefore conclude that the quiescence hypothesis is applicable to the Parkfield area¹⁹.

The present seismicity rates for earthquakes with $M > 2.4$ and $M > 2.0$ are compared to the background rate in Fig. 3. It is obvious that the rate has decreased by 80 and 46%, respectively. This quiescence is present in all of the Parkfield segment (Fig. 2) except in the northernmost few kilometres. The onset of the

quiescence is September 1986 for $M > 2.0$ events (for $M > 2.4$ it is January 1986) so that, at the time of writing, it has lasted ~ 3.5 years.

How should this quiescence anomaly be interpreted? The magnitude signature indicates that it is not artificial and the workers producing the catalogue, who also noticed the quiescence²⁰, do not think that the observed rate decrease in events of relatively large magnitude can be artificial. We conclude that the current Parkfield quiescence is probably natural and may be interpreted using the quiescence hypothesis. Based on the criteria from the Stone Canyon case we judge the current Parkfield quiescence significant. In the north the anomaly seems to be terminated near latitude 36.05°N . The southern extent of the anomaly cannot be well defined because the number of earthquakes per volume decreases towards the south to near zero at the southern end of the polygon shown in Fig. 2.

All aspects of the observed anomaly agree with the quiescence hypothesis, except that it is observed only for earthquakes of $M > 2$ and not for smaller events. The long duration of the anomaly, and its large spatial extent (~ 40 km along the fault) strongly suggest that this anomaly is part of the preparatory process for a fairly large rupture. We therefore conclude that the observed quiescence heralds the expected Parkfield earthquake ($M_L = 5.7$), not a smaller mainshock.

The time window for the expected earthquake was ten years^{1,2}, almost 50% of the recurrence period. We believe we can reduce this. In a previous test of the quiescence hypothesis in real time, we had predicted that two moderate earthquakes would rupture the San Andreas fault along specific segments²¹. One of these predictions was incorrect, the other was fulfilled by a rupture of the exact fault segment specified⁴. In that case the anomalous fault segment was a factor of four shorter than at Parkfield. At the time when the prediction was made the anomaly had lasted 1.9 years and the uncertainty of the duration was estimated to be ± 0.5 yr (ref. 21).

In the Aleutian islands a large earthquake was predicted based on quiescence^{5,6} that had lasted 3.5 years and which had been terminated by an $M_s = 8$ mainshock (M_s is the surface wave magnitude). In Hawaii $M_s = 6.6$ and $M_s = 7.2$ earthquakes were preceded by precursory quiescences of 2.4 and 3.8 years^{10,11}. Comparing the present quiescence of 3.5 years at Parkfield with these values indicates that the Parkfield earthquake must be

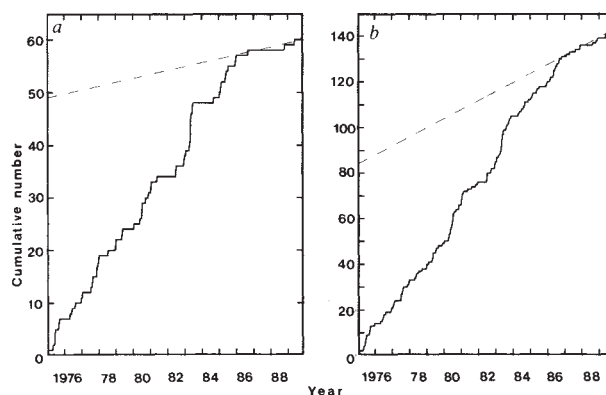


FIG. 3 Cumulative number of earthquakes as a function of time for the Parkfield area outlined by a rectangle in Fig. 2 for a, $M > 2.4$ and b, $M > 2.0$. The current, anomalously low seismicity rate is extrapolated back in time by a dashed line for comparison with the much larger observed background rate, represented by the slope of the solid curve. Because this pattern is very similar to that observed before other mainshocks along the San Andreas fault (Fig. 1) we interpret it to indicate that the next Parkfield earthquake will occur within the next 2 years. The rates decreased by 80% and 40% respectively for $M > 2.4$ and $M > 2.0$ earthquakes.

imminent. Along the San Andreas fault (Stone Canyon and Parkfield), however, quiescences of 1.5, 2.9 and 1.5 years were observed before mainshocks with magnitudes of 5.0, 4.7 and 3.6, respectively. It seems that along this segment of the San Andreas fault quiescences last longer than at other locations for mainshocks of the same magnitude.

Many workers have suggested that the duration of precursory quiescence should increase with increasing magnitude. The data suggest that this may be so, but are not conclusive⁸. In Hawaii an increase of one order in magnitude increases the duration of quiescence by about 1.5 years. The next characteristic Parkfield earthquake is expected^{1,2} to have a magnitude about two units larger than the prototype event in Parkfield (Fig. 1b, $M_L = 3.6$). Thus one may propose that the quiescence before the next characteristic earthquake should be about 3 years longer than the quiescence before the $M_L = 3.6$ event. This would lead to an expected precursor duration for the Parkfield earthquake of ~4.5 years. It is not possible to calculate rigorously an error for the expected precursor time because few data points are available. From the variations of precursor times for earthquakes along the San Andreas fault and the differences in their magnitudes, we estimate that the uncertainty in the expected precursor time may be about one year.

We feel that the above inferences are weak, but this is unavoidable in the absence of more data. We therefore estimate the occurrence time of the next Parkfield earthquake as March 1991 ± 1 year (4.5 years after September 1986, the average onset time of the anomaly). The magnitude and location of this event were given by Bakun and McEvilly¹.

The estimate of the probability with which the characteristic Parkfield earthquake is expected within the next 2 years (our range), without the information presented here, depends on the assumptions made. If the six known Parkfield mainshocks since 1850 were all characteristic events and are assumed to occur at random times, there is an 8% chance for the next one in any 2-year period. But if one uses the method of Nishenko and Buland²² or of Davis *et al.*²³ to analyse the recurrence time, the estimate is 38 and 14%, respectively. Based on our systematic searches for quiescent periods that do not precede mainshocks, we estimate that there is a less than 50% chance that the quiescence at Parkfield is a false alarm. Thus, the probability gain owing to seismic quiescence ranges from a factor of 1.3 to 3, depending on the assumptions.

The time window of uncertainty proposed by us is two years, which is one-fifth of the original ten-year window but, because only four years of the original window are left, this time is only shortened by 50%. Thus the information contained in the decrease of the seismicity rate is useful, even though the quiescence method is in its infancy. The fact that our time window falls within the original window is encouraging. As the two methods of estimating the occurrence time of the next Parkfield earthquake are based on independent data and independent models, the agreement in their results represents strong mutual support, and increased confidence that the Parkfield earthquake will occur soon. Other supporting evidence comes from a decrease in shortening rate for geodetic lines crossing the fault north of Parkfield in mid-1986. This observation, which will be reported in detail elsewhere²⁴, also indicates that a change in the faulting process took place in mid-1986. We interpret these two phenomena as intermediate-term precursors. Other intermediate-term precursors, if they exist in other data sets should also show up in mid-1986, and short-term precursors may still occur in the future, allowing a further reduction in the time window of the prediction of the next Parkfield earthquake. $\square$

5. Kisslinger, C., McDonald, C. & Bowman, J. R. *Proc. IASPEI* 32 (1985).
6. Kisslinger, C. *Bull. seism. Soc. Am.* **78**, 218–229 (1988).
7. Reasenber, P. *J. geophys. Res.* **90**, 5479–5495 (1985).
8. Wyss, M. & Habermann, R. E. *Pure appl. Geophys.* **126**, 319–332 (1988).
9. Wyss, M. & Habermann, R. E. *Pure appl. Geophys.* **126**, 333–356 (1988).
10. Wyss, M. *Bull. seism. Soc. Am.* **76**, 785–800 (1986).
11. Wyss, M., Klein, F. W. & Johnston, A. C. *J. geophys. Res.* **86**, 3881–3900 (1981).
12. Cao, T. & Aki, K. *Geol. Surv. U.S., Open file Rep.* 87-591 934–957 (1987).
13. Snoke, J. A., Taylor, D. W. A. & Sacks, I. S. *Eos* **70**, 1229 (1989).
14. Habermann, R. E. *J. geophys. Res.* **88**, 5056–5064 (1983).
15. Habermann, R. E. *Pure appl. Geophys.* **126**, 279–318 (1988).
16. Habermann, R. E. *Bull. seism. Soc. Am.* **72**, 93–112 (1982).
17. Habermann, R. E. *Bull. seism. Soc. Am.* **76**, 1660–1667 (1986).
18. Habermann, R. E. & Wyss, M. *Bull. seism. Soc. Am.* **74**, 1743–1755 (1984).
19. Bodin, P., Wyss, M. & Habermann, R. E. *Eos* **70**, 1230 (1989).
20. Aviles, C. A. & Valdes, C. M. *Eos* **70**, 1228 (1989).
21. Wyss, M. & Burford, R. O. *Geol. Surv. U.S. Open File Rep.* 85-745 367–426 (1985).
22. Nishenko, S. P. & Buland, R. *Bull. seism. Soc. Am.* **77**, 1382–1399 (1987).
23. Davis, P. M., Jackson, D. D. & Kagan, Y. Y. *Bull. seism. Soc. Am.* **79**, 1439–1450 (1989).
24. Wyss, M., Slater, L. & Burford, R. O. *Nature* **345**, 428–431 (1990).

ACKNOWLEDGEMENTS. We thank C. Aviles for discussions, and R. Bilham and C. Kisslinger for reading the manuscript. We thank the US Geological Survey staff for making available the Parkfield earthquake catalogue data. This work was partially supported by the US Geological Survey.

Decrease in deformation rate as a possible precursor to the next Parkfield earthquake

M. Wyss*, L. Slater* & R. O. Burford†

* CIRES/Department of Geology, University of Colorado, Boulder, Colorado 80309, USA

† US Geological Survey, Menlo Park, California 94025, USA

CRUSTAL deformations along fault zones are monitored because they show the accumulation and possible premonitory release of strain energy before major earthquakes. On the Parkfield segment of the San Andreas fault the next mainshock (local magnitude $M_L = 5.7$) is expected to occur between 1983 and 1993, based on the relatively regular recurrence of characteristic earthquakes in 1857, 1881, 1901, 1922, 1934 and 1966¹. A concerted monitoring effort is therefore underway at Parkfield, with the aim of detecting precursory changes in observable quantities that can be measured reliably². Here we report data from two geodetic lines crossing the San Andreas fault near Parkfield which show that the rate of shortening decreased by ~20% starting in August 1986. This change coincided with a rate decrease of 45% for earthquakes of $M_L \geq 2.1$ in the area^{3,4}. Both phenomena may have been caused by temporary locking of the fault surface, or a decrease in strain accumulation rate, and they are interpreted as intermediate-term precursors to the next Parkfield earthquake.

The monitoring of crustal deformations near the fault using a two-colour laser distance measuring instrument was begun in 1984. The accuracy of the measurements is 10^{-7} times the line length^{5,6} thus the repeatability for a 5-km-long line is 0.5 mm. The instrument, located near the centre of the 1966 rupture zone (Fig. 2), simultaneously transmits two laser beams (wavelengths 441.6 nm and 632.8 nm) to reflectors that return them to the observation site, where the two-way travel times are measured. The slight difference in travel times between the two beams allows highly accurate corrections for changing atmospheric conditions⁵.

The lengths of the lines have been measured on about 700 days since mid-1984 (almost three times per week). Seasonal fluctuations, observed for most line lengths⁷, can be as large as 5 mm for the line most strongly affected. For least-squares straight-line fits to the data from lines not strongly affected by seasonal changes, the standard deviations range from 0.06 to 0.2 mm yr⁻¹. We should therefore be able to resolve changes in deformation rate of less than 1 mm yr⁻¹.

The length changes for lines crossing the San Andreas fault (Fig. 2) were examined carefully (Fig. 1a) after it was discovered

Received 7 February; accepted 6 April 1990.

1. Bakun, W. H. & McEvilly, T. V. *J. geophys. Res.* **89**, 3051–3058 (1984).
2. Bakun, W. H. & Lindh, A. G. *Science* **229**, 619–624 (1985).
3. Brown, R. D. Jr *et al. Geol. Surv., Prof. Pap. U.S.* **597** (1967).
4. Wyss, M. & Burford, R. O. *Nature* **329**, 323–325 (1987).

that seismic quiescence had started in the Parkfield area in September 1986^{3,4,8}. This attempt to identify a geodetic precursor was motivated by the successful prediction of an earthquake ($M_L = 4.7$, May 1986) on the San Andreas fault 80 km northwest of Parkfield, based on simultaneous precursory seismic quiescence and reduction in fault-creep (aseismic slip along the fault surface)^{9,10}.

To test the hypothesis that the deformation rate near Parkfield may have been reduced synchronously with the seismic quies-

cence, we estimated the average shortening rate for early and late sections of the record by a least-squares linear fit of both. The comparison was made for divisions between the two sections at every measurement in Fig. 1a, starting with the eightieth point and ending with the eightieth-before-final data point. The division points that yielded the largest contrast between early and late sections were in August 1986 for all four data sets of Fig. 3. Thus the shortening rate decrease and the seismic quiescence started at the same time within the resolving power of the data, which is estimated to be no better than one month and several months for the seismicity and deformation rates, respectively. The decreases in rate are 1.72 and 1.67 mm yr^{-1} for CAN and BAR respectively, and the standard deviations for the straight-line fits before and after August 1986 are 0.15 and 0.05 mm yr^{-1} respectively for both lines. If we assume that displacements are parallel to the fault, the average decrease in deformation rate is estimated to be $\sim 2.1 \text{ mm yr}^{-1}$.

The difference in shortening rate before and after August 1986 at BAR and CAN (10.1 to 8.4 mm yr^{-1}) is significant because it is more than 10 times the standard deviations of the slope changes between the two segments of the data. We therefore feel that the data present a strong case for a real change in the relative displacement rates of BAR and CAN.

To reduce or eliminate possible errors (mistakes in meteorological measurements, instrument instability or electronic drift) we calculated the difference between the distances of two lines extending in similar directions but to benchmarks on opposite sides of the fault, measured at nearly the same time (within minutes). The differences BAR-LANG and CAN-LANG (Figs 1 and 3 respectively) show the rate change even more clearly. If we assume that changes in the difference between these lines is the result of relative displacement of the two benchmarks BAR and LANG parallel to the fault, we calculate that the differential displacement rate changed from 16.3 to 13.0 mm yr^{-1} in August 1986. The interruption of the time series

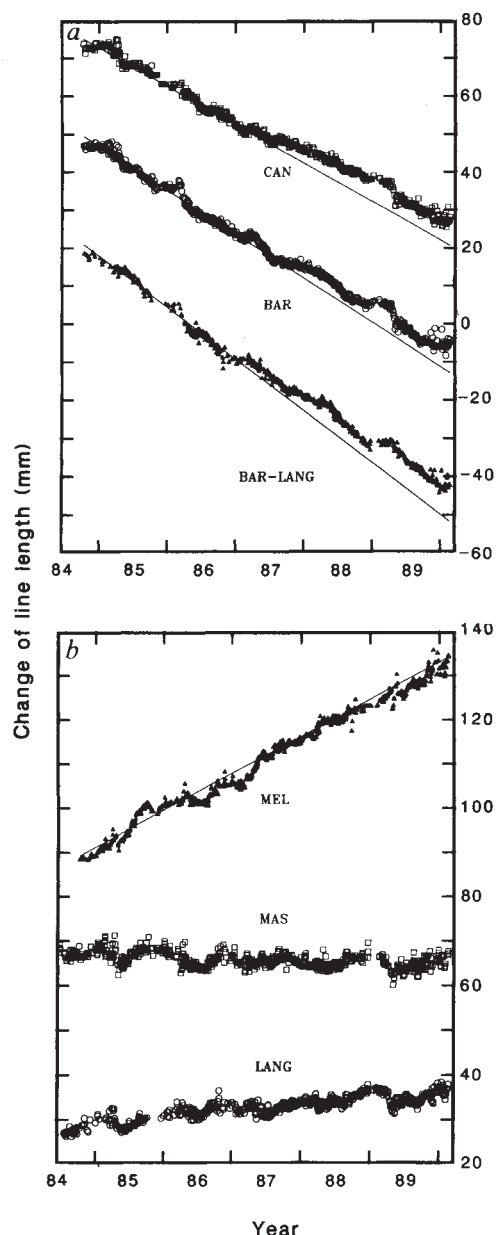


FIG. 1 a Change of length as a function of time for two lines crossing the San Andreas observed from CAR to benchmarks situated north of Parkfield. Squares and circles are for lines to CAN and BAR respectively (Fig. 2). Straight lines show the least-squares fits for the period between the start of the data and August 1986. The shortening of these lines indicates fault slip or strain accumulation with the eastern side of the fault moving to the southeast relative to the observation site at Car Hill. In the differences between lengths to BAR and LANG (triangles), errors caused by meteorological influence, instrument movement or electronic instabilities are removed to a great extent. b Change of length for two lines that do not cross the fault (LANG and MAS) and for one line that crosses the fault to the south (MEL).

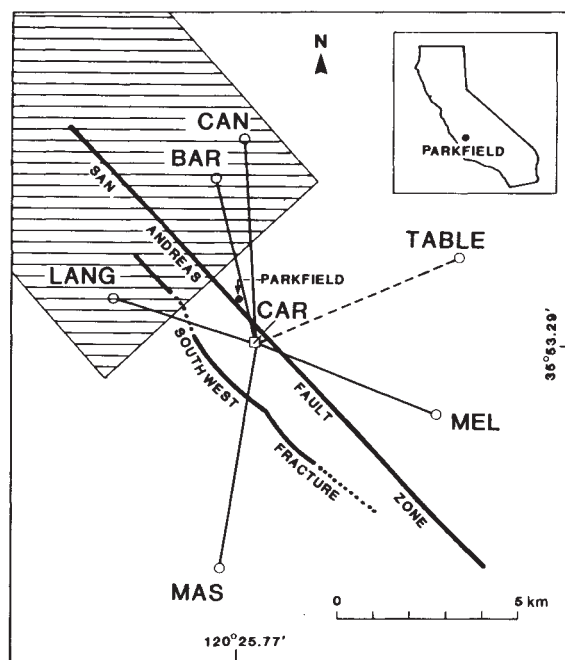


FIG. 2 Map of the Parkfield area with solid lines connecting the two-colour laser observation site CAR with the reflector points for which line-length data are shown in Fig. 1. The approximate location of the San Andreas fault is also indicated. The fault segment with the most pronounced seismic quiescence is hatched. The length of the dashed line was affected by the Kettleman Hills earthquake of 1985⁷.

for the baseline to CAN⁷ does not seem to be a source of error because the data for CAN and BAR give the same results.

The change of displacement rate is best seen in plots of the residuals with respect to the straight-line fits through August 1986 (Fig. 3). At present, a shortening deficiency of ~5 mm has accumulated (7 mm projected as displacement deficiency parallel to the fault). The accumulated differential displacement deficiency (BAR-LANG and CAN-LANG) is ~10 mm. We conclude from this data analysis that in the later part of 1986 a decrease in the displacement rate of ~20% took place on the San Andreas fault northwest of Parkfield.

We could not detect significant changes in rates of shortening or lengthening in other lines during 1986 and 1987. Examples of length changes for lines not crossing the fault to bench marks with small seasonal changes are shown in Fig. 1b. The fault-crossing line south of the observation station (CAR to MEL in Fig. 2) did not show a significant difference in the fits for the early and late sections. Thus we propose that the deformation rate change took place only northwest of Parkfield. This suggestion is also in agreement with the observations of the seismicity rate, which indicate that the seismic quiescence was most strongly developed in the fault segment northwest of Parkfield^{3,4} as indicated by the hatched rectangle in Fig. 2. The data for other lines⁷ are not discussed here because they either start too late or contain strong seasonal changes.

No moderate to large earthquakes took place in the vicinity of Parkfield within a year before and since the time when the observed decrease in displacement rate took place. Before the two-colour laser was installed, the Coalinga earthquake ($M_L = 6.9$, 2 May 1983, located ~45 km from Parkfield off the San Andreas fault) caused a perturbation in the creep observed along the Parkfield segment of the fault^{11,13}. The Kettleman Hills earthquake ($M_L = 5.5$, 4 August 1985, located ~35 km to the northeast of Parkfield) changed the laser line length to TABLE, but otherwise had only small effects on line lengths and creep records¹². Therefore we know that tectonic events at some distance can influence the slip rate along the Parkfield fault segment, but we see no external cause to which the August 1986 lasting rate decrease could be attributed. Also it is not likely that changing moisture content of the ground could have caused the

lasting rate change because the average annual amount of rainfall during the laser observation period was the same, within 10%, before and after August 1986.

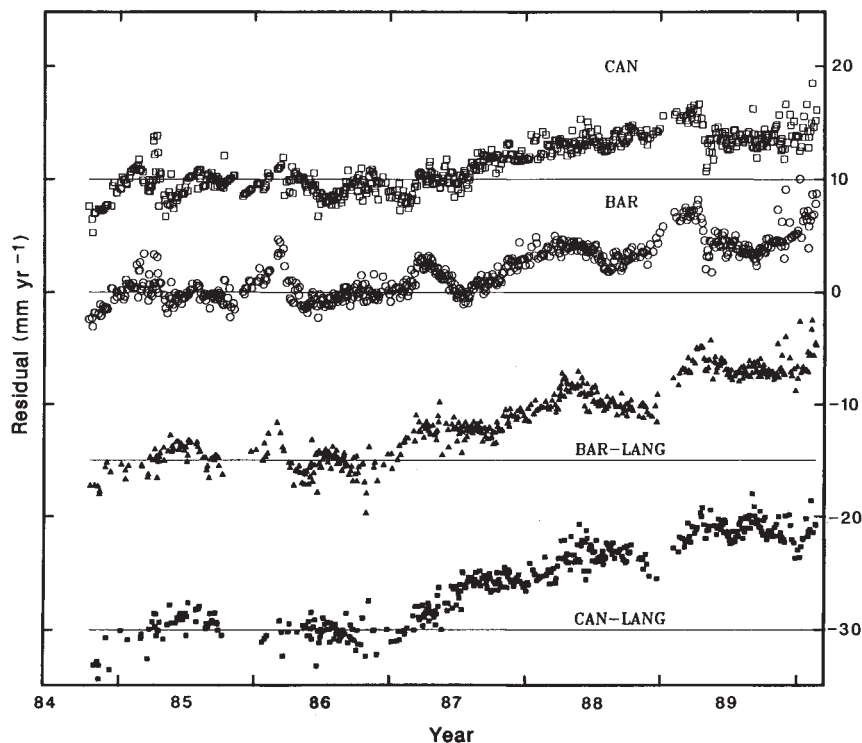
Aseismic slip measured at the surface by creepmeters does not show a decrease in rate at any time between 1984 and 1989 but there may have been an acceleration in early 1986 at a creepmeter south of Parkfield⁷. Langbein *et al.*⁷ modelled the observed deformations in the laser network by varying fault slip at depth as a function of space and time. Their Fig. 10 shows that, according to their model, a decrease of fault slip rate occurred north of Parkfield in 1986.

Partial locking of the fault¹³ could decrease both the rate of seismic and aseismic slip. But this mechanism would lead to an increase in ambient stress which should result in a decrease in b -value¹⁴ (equivalent to an increase in average magnitude of earthquakes produced), yet the opposite is observed: the b -value increased from 1.1 to 1.4 as we shall show elsewhere. If an increase in b -value is taken to signal a decrease in ambient shear stress, or an increase in pore pressure¹⁵, it seems difficult to explain the reduction in deformation rate, because the latter should increase if pore pressure is increased or if strain is released. The mechanism that causes seismic quiescence and deformation rate decrease cannot be ascertained yet because not enough constraints exist.

The seismic quiescence and the reduction in fault slip now in effect at Parkfield show patterns very similar to those observed before the successfully predicted Stone Canyon earthquake^{9,10}. We propose that these two phenomena strongly suggest that an intermediate-term preparatory process to a characteristic Parkfield earthquake¹ started in late 1986, and that this earthquake is to be expected soon. We have argued that, based on comparisons with other case histories of seismic quiescence, the Parkfield anomaly has already lasted for an unusually long period, therefore the mainshock is expected in the near future, and no later than March 1992 (ref. 4).

The occurrence of deformation rate decrease as an intermediate-term precursor does not preclude accelerated slip during weeks to days before the next mainshock as expected on the basis of some models of the failure process¹⁶. The possibility of a short-term increase in slip rate is suggested by the observa-

FIG. 3 Residuals of the rate of length changes of the lines to CAN (squares), BAR (circles) and the difference BAR-LANG (triangles), CAN-LANG (squares) with respect to the least-squares fits of data for the period between October 1984 and August 1986. Changes of line length are shown in Fig. 1 and locations of the lines are given in Fig. 2.



tion of relatively fresh en echelon cracks along the fault during the month before the last Parkfield mainshock¹⁷, and by the rupture of a water main only hours before it¹⁸. The relationship between intermediate-term precursors (months to years) and short-term precursors (hours to weeks) is unknown.

It is important that independent precursors are observed in order to gain confidence during the process of refining an earthquake prediction. The observations of seismic quiescence and deformation deficiency are independent and are affected by different sources of error. Thus their nearly synchronous onset strongly suggests that these lasting changes are real. The two types of changes are probably not physically independent, however, because a common underlying process may cause both phenomena.

It is also important to attempt the identification of precursors before the mainshock has occurred in order to learn what misinterpretations are made when one is faced with data which may contain potential precursors. This attempted earthquake prediction is notable in being based on four lines of evidence: regular recurrence, seismicity rate decrease of 45%, deformation rate decrease of 20% and a *b*-value increase of 25%. □

Received 15 March; accepted 18 April 1990.

1. Bakun, W. H. & McEvilly, T. V. *J. geophys. Res.* **89**, 3051–3058 (1984).
2. Bakun, W. H. & Lindh, A. G. *Science* **229**, 619–624 (1985).
3. Bodin, P., Wyss, M. & Habermann, R. E. *Eos* **70**, 1230 (1989).
4. Wyss, M., Bodin, P. & Habermann, R. E. *Nature* **345**, 426–428 (1990).
5. Slater, L. E. & Hugggett, G. R. *J. geophys. Res.* **81**, 6299–6306 (1976).
6. Burford, R. O. & Slater, L. E. *U.S. Geol. Survey, Open File Report*, 85–754, 117–129 (1985).
7. Langbein, L. E., Burford, R. O. & Slater, L. E. *J. geophys. Res.* **95**, 2533–2552 (1990).
8. Aviles, C. A. & Valdes, C. M. *Eos* **70**, 1228 (1989).
9. Wyss, M. & Burford, R. O. *U.S. Geol. Survey, Open File Report* 85–745, 367–426 (1985).
10. Wyss, M. & Burford, R. O. *Nature* **329**, 323–325 (1987).
11. Burford, R. O., Schulz, S. S. & Simpson, R. W. *U.S. Geol. Survey Open File Report* 87–591, 845–867 (1987).
12. Simpson, R. W., Schulz, S. S., Dietz, L. D. & Burford, R. O. *Pure appl. Geophys.* **126**, 665–685 (1988).
13. Burford, R. O. *Pure appl. Geophys.* **126**, 499–529 (1988).
14. Scholz, C. H. *Bull. seism. Soc. Am.* **58**, 399–416, (1968).
15. Wyss, M. *Geophys. J. R. astr. Soc.* **31**, 341–359 (1973).
16. Stuart, W. D. *J. geophys. Res.* **84**, 1063–1070 (1979).
17. Allen, C. R. & Smith, S. W. *Bull. seism. Soc. Am.* **56**, 966–967 (1966).
18. Brown, R. D. Jr et al. *Geol. Surv. Prof. Pap. U.S.* **597**, (1967).

ACKNOWLEDGMENTS. We thank W. D. Stuart for discussions and comments, and R. Bilham, E. Roeloffs and C. Morte for comments on the manuscript. This work was supported in part by the U.S. Geological Survey external grants program.

Biological recovery of an acid lake after reductions in industrial emissions of sulphur

J. M. Gunn* & W. Keller†

* Ontario Ministry of Natural Resources, Department of Biology, Laurentian University, Sudbury, Ontario P3E 2C6, Canada

† Ontario Ministry of The Environment, 199 Larch Street, Sudbury, Ontario P3E 5P9, Canada

THERE is now considerable evidence from Canada^{1–4}, Norway⁵, Sweden⁶ and Scotland⁷ that lake and stream acidification is chemically reversible. Water quality improvements have been shown to occur rapidly when inputs of sulphur are reduced⁸. However, there is little evidence that biological systems recover after emission reductions⁹. We report here on changes in the biological community of an acid-stressed lake near the metal smelters at Sudbury, Canada, during a 10-year period (1978–1988) after smelter emissions of sulphur were reduced by more than 50%. Our studies of the fish, benthic invertebrates, and zooplankton in this lake show that rapid recovery and reinvasion occurs among some acid-sensitive species.

The study was conducted on Whitepine Lake (47°17'N, 80°50'W), a clear, dilute (conductivity around 30 $\mu\text{mho cm}^{-1}$)

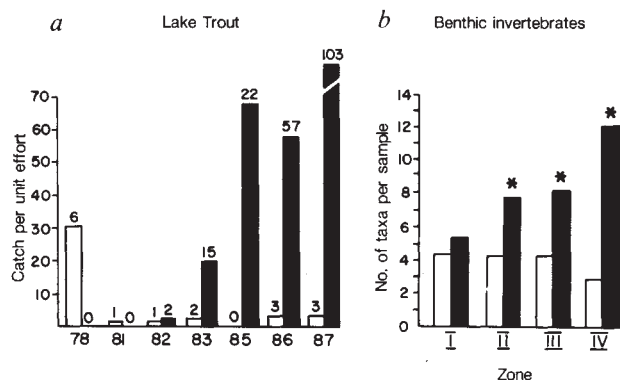


FIG. 1 *a*, Catch per unit of effort. Number of lake trout caught in experimental gillnets (1.8–7.6-cm mesh) in Whitepine Lake. Data standardized to 100 h of netting effort. Only naturally produced fish are included in the figure. The total number of adults (open columns) and juveniles (solid columns) captured in each year is indicated. *b*, Changes in the number of taxa of benthic invertebrates in Whitepine Lake. Three replicate samples were collected at each of three stations in each offshore zone (I, 8–21 m; II, 4–8 m; III, 1.5–4 m) using an 15 × 15 cm Ekman dredge on May 20, in both 1982 (open columns) and 1988 (solid columns). Samples were sieved through a 600- μm mesh screen before counting. Sampling was repeated in October of both years with similar results. The nearshore rocky areas (IV, <1 m) were sampled with artificial substrates (23 × 23 × 15 cm cages filled with 5–10-cm diameter clean stones) placed in triplicate at five sites. The placement and retrieval dates were 25 May and 19 July, respectively, in both 1983 (open columns) and 1988 (solid columns). Autumn samples were also collected using artificial samplers; again with similar results.

* Depth zones showing significant increases (*t*-test, $P \leq 0.05$) in number of taxa between sampling periods.

headwater lake (surface area 67 ha, maximum depth 22 m), located in a forested area 90 km north of Sudbury. The lake receives both wet and dry inputs of sulphur, but is outside the zone of high deposition of metal particulates from the Sudbury smelters¹ (water concentrations of Cu and Ni <2 $\mu\text{g l}^{-1}$). In 1980 the lake was acidic (pH <5.5) and its fish population showed classical signs of acid stress: acid-tolerant yellow perch (*Perca flavescens*) were very abundant, while acid-sensitive species such as lake trout (*Salvelinus namaycush*), burbot (*Lota lota*), and white suckers (*Catostomus commersoni*) were extremely rare and were not reproducing¹⁰.

Monitoring of lake water quality began in 1980. Fish populations were assessed during 1978, 1981, 1982, 1983, 1984, 1986, and 1987 using similar sampling sites and capture gear (variable mesh gillnets, trapnets, plexiglas minnow traps)¹¹. Benthic invertebrates inhabiting soft sediments were sampled in 1982 and 1988, with a depth-stratified sampling design and an Ekman dredge collector. Artificial substrate collectors were used in 1983 and 1988 to sample invertebrates that occupied rocky nearshore areas. Zooplankton were collected at least monthly during the open water periods of 1980–1988. Zooplankton samples were collected with a 12.5-cm diameter 80- μm mesh plankton net drawn vertically through the water column at a mid-lake site.

Water quality improved significantly ($P < 0.01$) from 1980 (pH 5.4, acid-neutralizing capacity (ANC) 1 $\mu\text{equivalent l}^{-1}$) to 1988 (pH 5.9, ANC 11 $\mu\text{equivalent l}^{-1}$) (Table 1). Specific conductance and concentrations of SO_4 , Ca and Al all declined with reductions in industrial emissions of SO_2 (Table 1).

The first evidence of biological recovery was the appearance of naturally produced young lake trout from the remnant adult population. Young lake trout first appeared in 1982 and became increasingly abundant throughout the rest of the study (Fig. 1). Production of young trout by the remnant adult population was discovered during routine monitoring of hatchery fish added to this lake in 1980 and 1981¹¹. Each hatchery fish was marked before addition by removing the adipose fin. During monitoring of these individuals an increasing number of unmarked fish

TABLE 1 Water quality data for Whitepine Lake, 1980–1988

	Sudbury emissions of SO ₂ (×10 ³ tonnes)*	pH	Acid-neutralizing capacity (μequivalent l ⁻¹)	Conductivity (μmho cm ⁻¹)	Ca (μequivalent l ⁻¹)	SO ₄ (μequivalent l ⁻¹)	Total Al (μg l ⁻¹)
1980	935	5.4 (5.3–5.8)	1 ± 5	37 ± 2	156 ± 18	237 ± 20	69 ± 20
1981	837	5.7 (5.6–5.8)	4 ± 2	35 ± 1	145 ± 31	235 ± 12	59 ± 21
1982	389	5.8 (5.8–5.9)	10 ± 4	36 ± 3	138 ± 17	218 ± 20	44 ± 15
1983	538	5.8 (5.7–5.8)	8 ± 3	32 ± 1	123 ± 5	198 ± 22	36 ± 12
1984	767	5.8 (5.7–5.8)	8 ± 2	32 ± 2	125 ± 10	198 ± 24	40 ± 15
1985	769	5.8 (5.7–5.9)	12 ± 4	31 ± 1	119 ± 13	195 ± 4	50 ± 11
1986	725	5.9 (5.8–5.9)	12 ± 4	29 ± 1	120 ± 9	177 ± 17	34 ± 10
1987	730	6.0 (5.9–6.1)	15 ± 1	30 ± 4	115 ± 13	196 ± 23	26 ± 11
1988	718	5.9 (5.9–6.0)	11 ± 3	30 ± 4	115 ± 6	195 ± 10	40 ± 14

Annual means (±1 s.d., range given for pH) are calculated from monthly data ($n=4-11$) for samples collected through the epilimnion and metalimnion. Changes through time were determined to be significant ($P < 0.01$) for all of the above chemical parameters using linear correlation analysis of log-transformed (except pH) data against time in months.

* Annual averages for SO₂ emissions (×10³ tonnes) before the study were: 1960–64, 2,241 tonnes; 1965–69, 2,202 tonnes; 1970–74, 1,867 tonnes; 1975–79, 1,064 tonnes.

occurred in yearly catches. By 1987, the naturally produced fish were far more abundant than the stocked fish, and dominated (87%) the gillnet catches. These young fish began to appear at least two years before the stocked fish had reached sexual maturity (age 5, Autumn 1984).

Other researchers have shown that lake trout do not reproduce when pH drops below 5.5–5.6 (refs 12–13). Our study showed that this effect is reversible; successful reproduction occurred when pH increased above 5.5.

Successful reproduction did not occur in the white sucker and burbot populations, and these species were not observed after 1984. Very few adults (less than five fish observed) were present for each species, a fact that presumably contributed to their extinction. In a lake where adult white suckers were abundant and were not reproducing under low pH conditions, successful reproduction has been shown to occur with an increase in pH (ref. 14).

The rapid increase in abundance of lake trout, an omnivorous predator, caused marked changes in prey populations. Trapnet catches of perch, a common prey species of lake trout¹⁵, declined

from 145 per night in 1982 to approximately 2 per night in 1987. Densities of benthic invertebrates in the deep zones of the lake (8–21 m) that are occupied by lake trout declined from 1140 m⁻² in 1982 to 650 m⁻² in 1988. In contrast, the density of invertebrates in the shallow areas (0–8 m), the principal foraging area of perch, increased with the decline in abundance of perch. Soft sediment areas provided 550 invertebrates m⁻² in 1982, 1,320 m⁻² in 1988; rocky areas, 6 invertebrates per sample in 1983, 164 per sample in 1988.

Although changes in fish predation may explain many of the changes in invertebrate populations, direct effects of water quality improvements were also detected. The number of taxa of benthic invertebrates increased from 39 in 1982/83 to 72 in 1988 (Fig. 1) and the relative abundance of many of the original species changed dramatically. Three new species of mayfly (*Hexagenia* sp., *Ephemerella* sp., *Caenis* sp.) and two species of oligochaetes (*Vejdovskyella comata*, *Slavina appendiculata*) appeared in the most recent survey. A single specimen of gastropoda (*Ferrissia* sp.) was also captured for the first time in 1988. Taxa scarce in the early survey, including leeches and crayfish (*Orconectes propinquus*), increased in numbers coincident with improvements in water quality. The benthic survey also indicated an increase in Iowa darters (*Etheostoma exile*), a bottom-dwelling fish. One new zooplankton species (*Epischura lacustris*) also appeared in the collections after 1981. On the basis of previous studies all the above species can be considered acid-sensitive^{16–22}.

Whitepine Lake is only one of many lakes in the Sudbury area where water quality has improved following reduced emissions of sulphur (Fig. 2). In January 1980 and January 1987, we sampled 66.6% ($n=104$) of all known lake trout lakes (ranging from extinct to healthy populations) within a 1.4×10^3 km² area around Sudbury. The sample set was slightly biased towards larger lakes (surface area 38–13,130 ha) to permit sampling by fixed-wing aircraft. Increases in pH ($\bar{x}\Delta\text{pH}=0.37$) and ANC ($\bar{x}\Delta\text{ANC}=23$ μequivalent l⁻¹) occurred in 98.1 and 97.3% of the lakes, respectively. The number of lakes with pH below the threshold for reproduction of lake trout (pH < 5.5) declined from 45 (43.3%) to 20 (19.2%) (Fig. 2).

In summary, our data show that rapid chemical and biological recovery of industrially-acidified lakes can be accomplished simply by reducing the emission to the atmosphere of acidifying substances. Conditions suitable for the survival and reproduction of lake trout and other acid-sensitive organisms can return without liming or other ameliorative treatments. □

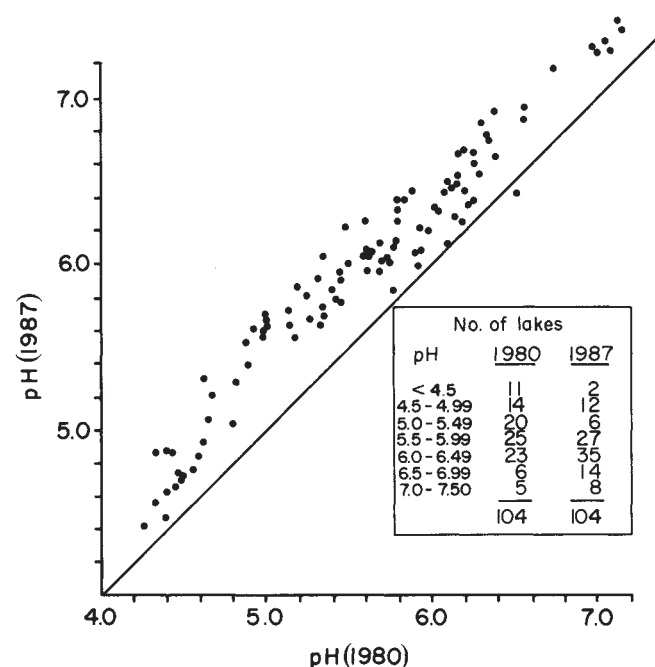


FIG. 2 Change in pH of 104 lake-trout lakes near Sudbury, Canada between 1980 and 1987. Water samples were collected with a 5-m tube composite sampler through the ice. The pH distribution of the sampled lakes is presented in the inset table. For further evidence of regional scale recovery of water quality see refs 1 and 2.

Received 2 October 1989; accepted 14 March 1990.

- Keller, W. & Pitblado, J. R. *Water Air Soil Pollut.* **29**, 285–296 (1986).
- Dillon, P. J., Reid, R. A. & Girard, R. *Water Air Soil Pollut.* **31**, 59–66 (1986).
- Thompson, M. E. *Water Air Soil Pollut.* **31**, 17–26 (1986).
- Hutchinson, T. C. & Havas, M. *Water Air Soil Pollut.* **28**, 319–333 (1986).
- Wright, R. F., Lotse, E. & Semb, A. *Nature* **334**, 670–675 (1988).

6. Forsberg, C., Morling, G. & Wetzel, R. G. *Ambio* **14**, 164–166 (1985).
7. Battarbee, R. W. et al. *Nature* **332**, 530–532 (1988).
8. Haughe, M. & Wright, R. F. *Acid Deposition: Reversibility of Soil and Water Acidification, A Review*, Air Pollution Report No. 11 (Commission of the European Communities, Brussels, 1988).
9. Schindler, D. W. *Science* **239**, 149–157 (1988).
10. Gunn, J. M. & Keller, W. *Can. J. Fish. Aquat. Sci.* **41**, 319–329 (1984).
11. Gunn, J. M., McMurtry, M. J., Bowly, J. N., Casselman, J. M. & Llimatainen, V. A. *Trans. Amer. Fish. Soc.* **116**, 618–627 (1987).
12. Beggs, G. L. & Gunn, J. M. *Water Air Soil Pollut.* **30**, 711–717 (1986).
13. Mills, K. H., Chalanchuk, S. M., Mohr, L. C. & Davies, I. J. *Can. J. Fish. Aquat. Sci.* **44** (Suppl. 1), 114–125 (1987).
14. Schindler, D. W. *Reversibility of Acidification* (ed. Barth, E.) 2–13 (Elsevier Applied Science, London, 1987).
15. Gunn, J. M. et al. *Can. J. Fish. Aquat. Sci.* **47**, 446–453 (1990).
16. Dermott, R. M., Kelso, J. R. M. & Douglas, A. *Water Air Soil Pollut.* **28**, 283–292 (1986).
17. Raddum, G. in *Ecological Impact of Acid Precipitation* (eds Drablos, D. & Tollan, A.) 330–331 (SNSF project, 1432 As, Norway, 1980).
18. Dermott, R. M. *Hydrobiologia* **128**, 31–38 (1985).
19. Berrill, M., Hollett, L., Margosian, A. & Hudson, J. *Can. J. Zool.* **63**, 2586–2589 (1985).
20. Rahel, F. J. & Magnuson, J. J. *Can. J. Fish. Aquat. Sci.* **40**, 3–9 (1983).
21. Keller, W. & Pitblado, J. R. *Water Air Soil Pollut.* **23**, 271–291 (1984).
22. Mills, K. H. & Schindler, D. W. *Water Air Soil Pollut.* **30**, 779–789 (1986).

ACKNOWLEDGEMENTS. We thank D. Farara, P. Gale and V. Llimatainen for assistance and P. Dillon, D. Schindler, K. Mills and D. Wales for constructive criticism. Funding was provided by Ontario Ministry of Natural Resources and Ontario Ministry of the Environment through the APIOS program.

Tritrophic effects of a simple architectural mutation in pea plants

Peter Kareiva & Robert Sahakian

Department of Zoology, University of Washington, Seattle, Washington 98195, USA

WHEN studying interactions between trophic levels, ecologists often restrict their attention to two levels as a means of simplifying the analysis; unfortunately, this simplification can be misleading if tritrophic interactions (such as plant–herbivore–predator) cannot be understood by simply adding together pairwise interactions (plant–herbivore plus herbivore–predator, for example)^{1,2}. We examined the significance of tritrophic interactions by asking how the morphology of the common pea (*Pisum sativum*) influences the population growth of pea aphids (*Acyrtosiphon pisum*) in the presence and in the absence of a third trophic level. We found significant interactions between the first trophic level (peas) and the third trophic level (ladybird beetles) in determining aphid population growth. Our results point out how simple genetic changes can yield morphological variants in plants that differ dramatically in their resistance to herbivores due to the effects of plant architecture on enemies of the herbivore.

Numerous studies suggest that plant morphology may indirectly influence herbivores by affecting predators or parasites^{3–6}. To show that such phenomena represent tritrophic interactions, however, one needs to factor out the direct effects of plant morphology on herbivores from the indirect effects mediated by the third trophic level. We distinguished these direct and indirect effects by using a factorial design in which two different treatments of plant architecture (normal versus leafless peas) were crossed with two predator treatments (with and without ladybirds), and rates of aphid population growth were used as the response variable. A significant interaction between the predation and the plant architecture treatments would be evidence for a tritrophic interaction. Because we were also interested in the extent to which within-species variation in plant morphology shapes plant–herbivore–predator interactions, we sought architectural variants that were known to differ genetically at a limited number of loci. Previous investigations of the interplay between plant morphology and herbivore enemies have involved different species of plants^{3,7,8}, or different artificial substrates⁹, or varieties of the same species that differ at an unknown number of unspecified genetic loci^{4,5}. In contrast, we used highly backcrossed strains of peas that differed at only two loci: a normal variety and a leafless variety that is distinguished

by homozygous recessive mutations at the *afla* and the reduced stipule loci¹⁰.

For each experiment, individual pea plants were grown in pots 21 cm in diameter and 21 cm deep, and were enclosed in cages 32 cm tall. An excess of late instar (IV) or adult apterous aphids were first placed on each plant, and then aphids were removed until we achieved the desired initial density, which was the same for all plants within any experiment. Although we were not able to produce perfectly identical age distributions of aphids on all plants, we did not introduce any consistent biases. After the aphid populations were established, the desired number of predators (either *Coccinella septempunctata*, *Hippodamia variegata*, or *Propylea quatuordecimpunctata*) were added to plants in the 'with predator' treatment and all of the plants were caged.

Experiments were initiated with identically aged pea plants (3–5 weeks old, depending on the experiment) and were allowed to run for 2–4 days. Although 2–4 days might seem a short time over which to observe population change, it proved to be an interval over which aphid numbers changed by as much as $\pm 100\%$ (this is because our final census included newborn aphids).

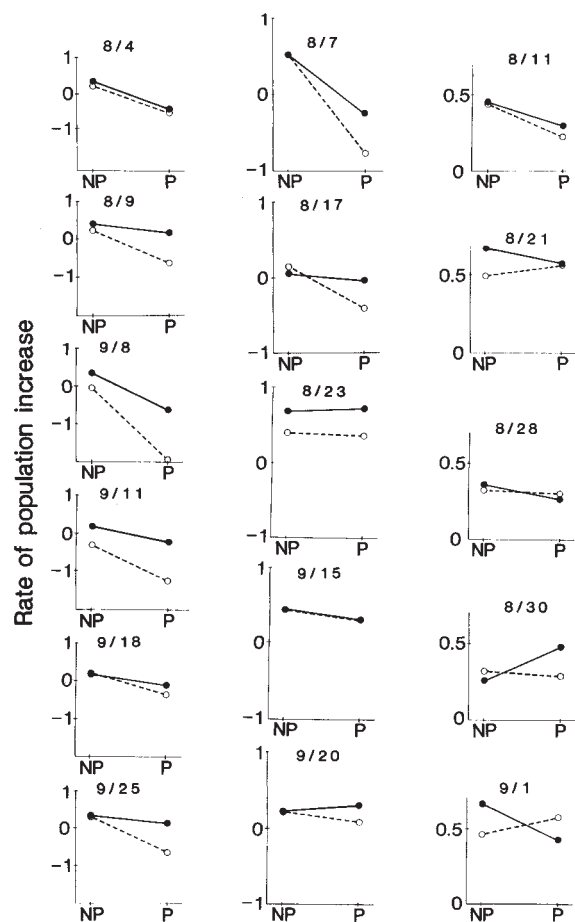


FIG. 1. Instantaneous rate of aphid population growth (per day) in factorial experiments that manipulated presence of ladybird predators (NP, no predators; P, with predators) and the architectural variety of pea. ---○---, Leafless mutant; —●—, normal leafy genotype. Left column of graphs, experiments with *Coccinella septempunctata*; middle column, *Hippodamia variegata*; right column, *Propylea quatuordecimpunctata*. Each set of axes is also labelled with the date on which the experiment terminated. The decline in aphid population growth from NP to P is a measure of predator impact; the differences in slope of dashed and solid lines reflect possible interaction effects (see Table 1 for statistical significance) or 'three-trophic-level effects'. The scale is expanded fourfold for *Propylea* versus *Coccinella* and *Hippodamia* columns; thus the effects of *Propylea* are negligible in comparison with *Coccinella* or *Hippodamia*.

TABLE 1 Summary of two-way analyses of variance using SYSTAT MGLH (ref. 17) with balanced designs and instantaneous rate of aphid population growth as the response variable

Predator	Architecture effects	Predation effects	Interaction effect	Aphid:beetle†
<i>Coccinella</i>	NS	**	NS	10:2
<i>Coccinella</i>	**	**	**	30:2
<i>Coccinella</i>	**	**	**	40:3
<i>Coccinella</i>	**	**	NS	40:2
<i>Coccinella</i>	*	**	*	40:2
<i>Coccinella</i>	**	**	*	30:2
<i>Hippodamia</i>	**	**	**	10:2
<i>Hippodamia</i>	**	**	**	20:2
<i>Hippodamia</i>	**	NS	NS	30:2
<i>Hippodamia</i>	NS	NS	NS	20:2
<i>Hippodamia</i>	NS	NS	*	40:3
<i>Propylea</i>	NS	**	NS	20:2
<i>Propylea</i>	NS	NS	NS	20:4
<i>Propylea</i>	NS	NS	NS	30:4
<i>Propylea</i>	NS	NS	NS	30:7
<i>Propylea</i>	NS	NS	*	30:7

Each row represents an independent experiment, using different sets of plants, aphids and ladybird beetles. NS, not significant.

* $P < 0.05$; ** $P < 0.01$.

† Number of aphids and coccinellids initially placed on plants.

TABLE 2 The effect of pea morphology on the frequency with which foraging ladybird beetles fall off pea plants while moving in search of aphids

Ladybird species	Leafless peas	Normal peas	
<i>Coccinella septempunctata</i>	26 of 100 (26%)	47 of 100 (47%)	$P < 0.01$
<i>Hippodamia variegata</i>	8 of 88 (9%)	28 of 88 (32%)	$P < 0.01$

One hundred (*Coccinella*) or 88 (*Hippodamia*) individual beetles were observed for 2 min on first one pea genotype and then the other pea genotype (first genotype determined randomly). Each beetle was used only once on the two pea genotypes, which justifies analysing the results as independent trials. Observations were spread over two weeks (7–23 July) in 1988 and took place on peas planted in the ground as opposed to pots. We used a log likelihood G-test to determine whether the fraction of ladybird beetles that fell off normal peas was different from the fraction that fell off leafless peas during the brief 2-min observations.

To analyse our results we focused on the instantaneous rate of aphid population growth defined by $\ln(N_t/N_0)/t$, where N_t is the final number of aphids on the pea plant, N_0 is the initial number of aphids on the pea plant, and t is the number of days for which the experiment had been run. Using a two-way analysis of variance we asked whether there were significant effects due to the predator treatments, the plant architecture treatments, or an interaction between predation and architecture treatments. When both main effects were significant, there was generally a significant interaction between the predation and architecture treatments (in four of six occasions, Table 1). If we discount *Propylea*, which had difficulty capturing and eating pea aphids, there was a striking trend for predators to be more effective at controlling aphid populations on leafless as opposed to normal peas (in 9 of the 11 comparisons involving either *Hippodamia* or *Coccinella*, Fig. 1). Aphids better escape predation by coccinellids on normal compared with leafless peas simply because ladybirds fall off the normal variety nearly twice as frequently as they do the leafless variety (Table 2); direct observation suggests this is because the beetles can cling to stems and tendrils but have difficulty maintaining a secure grasp on smooth, slippery leaves.

Our experiments show that we may be unable to detect the importance of plant morphological variation with respect to herbivory unless a third trophic level is included in the analysis. In the absence of a third trophic level, we found that aphid population growth was only slightly reduced (on average less

than 10%) on a mutant variety of peas compared with the normal genotype; however, in the presence of a third trophic level, a simple mutation to the leafless form produced peas that enjoyed greatly reduced rates of aphid population growth (on average greater than a 50% reduction) when compared with the normal genotype. Field observations using these same varieties of peas and suite of predators indicate that the interaction between plant architecture and a third trophic level is even more striking when the ladybirds roam freely from plant to plant as opposed to being confined in a cage¹². Since plant morphology often influences predator and parasite searching patterns^{3–5,7,8,11,13–15}, consideration of the third trophic level needs to be incorporated into breeding programs aimed at producing resistant morphological varieties of crops⁵. Moreover, in natural plant-insect associations, there is clearly the opportunity for subtle changes in plant architecture dramatically to alter the outcome of plant-herbivore interactions because of effects at the third trophic level. □

Received 18 January; accepted 20 March 1990.

1. Lawton, J. & McNeill, S. *Symp. Brit. Ecol. Soc.* **20**, 223–244 (1979).
2. Price, P. et al. *A. Rev. Ecol. Syst.* **76**, 456–462 (1983).
3. Hulsas-Jordan, P. & Van Lenteren, J. *Med. Fac. Landbouw. Rijksuniv. Gent* **43**, 431–440 (1978).
4. Kauffman, W. C. & Kennedy, G. G. *Envir. Ent.* **18**, 698–704 (1989).
5. Obrycki, J. J., Tauber, M. J. & Tingey, W. M. *J. econ. Ent.* **76**, 456–462 (1983).
6. Southwood, T. R. E. in *Insects and the Plant Surface* (eds Juniper, B. & Southwood, T. R. E.) 1–22 (Edward Arnold, London, 1986).
7. Carter, M. C., Sutherland, D. & Dixon, A. F. G. *Oecologia* **63**, 394–397 (1984).
8. Evans, E. W. *N.Y. ent. Soc.* **XC**, 129–133 (1982).
9. Andow, D. L. & Prokrym, D. R. *Oecologia* **82**, 162–165 (1990).
10. Wehner, T. C. & Gritton, E. T. *J. Am. Soc. hort. Sci.* **106**, 272–278 (1981).
11. Belcher, D. W. & Thurston, R. *Envir. Ent.* **11**, 91–94 (1982).
12. Kareiva, P. in *Critical Issues in Biological Control* (eds Mackauer, E., Ehler, L. & Roland, J.) 213–238 (Intercept, Andover (UK), 1990).
13. Quilici, S. & Iperiti, G. in *Ecology of Aphidophaga* (ed. Hodek, I.) 99–106 (Academia, Prague/Dr W. Junk, Dordrecht, 1986).
14. Rotheray, G. E. in *Ecology of Aphidophaga* (ed. Hodek, I.) 107–111 (Academia, Prague/Dr W. Junk, Dordrecht, 1986).
15. Shahjahan, M. & Streams, F. A. *Envir. Ent.* **2**, 921–925 (1973).
16. Bernays, E. & Graham, M. *Ecology* **69**, 886–892 (1988).
17. Wilkinson, L. *Systat, the System for Statistics* (SYSTAT Inc., Evanston, 1989).

ACKNOWLEDGEMENTS. This research was supported by a grant from NSF. The USDA APHIS laboratory in Niles, Michigan, provided technical support. E. T. Gritton (University of Wisconsin) provided pea seeds. We thank M. Peterson and R. B. Root for editorial comments.

Horizontal and vertical components of head movement are controlled by distinct neural circuits in the barn owl

Tom Masino & Eric I. Knudsen

Department of Neurobiology, Fairchild Science Building, Stanford University School of Medicine, Stanford, California 94305-5401, USA

To generate behaviour, the brain must transform sensory information into signals that are appropriate to control movement. Sensory and motor coordinate frames are fundamentally different, however: sensory coordinates are based on the spatiotemporal patterns of activity arising from the various sense organs, whereas motor coordinates are based on the pulling directions of muscles or groups of muscles. Results from psychophysical experiments suggest that in the process of transforming sensory information into motor control signals, the brain encodes movements in abstract or extrinsic coordinate frames^{1–5}, that is ones not closely related to the geometry of the sensory apparatus or of the skeletomusculature. Here we show that an abstract code underlies movements of the head by the barn owl. Specifically, the data show that subsequent to the retinotopic code for space in the optic tectum yet before the motor neuron code for muscle tensions there exists a code for head movement in which upward, downward, leftward

and rightward components of movement are controlled by four functionally distinct neural circuits. Such independent coding of orthogonal components of movement may be a common intermediate step in the transformation of sensation into behaviour.

Electrical microstimulation of the owl's optic tectum elicits a quick (saccadic), orienting movement of the head. These movements are similar to those evoked by tectal stimulation in other species⁶⁻¹⁰ in that the magnitude and direction (vector) of the movement varies systematically with the site of stimulation revealing a topographic code for movement vector¹¹. In analysing the transformation of the tectal motor code into neck motor neuron control signals, we exploited an experimental paradigm, developed by Robinson¹², that takes advantage of the fact that the post-tectal circuitry enters a brief refractory state after being activated. This refractory phenomenon is most clearly demonstrated when a pair of stimulus pulse trains is applied to a single site in the tectum. Two short (40 ms) pulse trains separated by more than 200 ms (interval between train onsets) elicit two similar movements of the head each lasting about 50 ms. But when these pulse trains are delivered at shorter intervals, the movement in response to the second pulse train is reduced and, at intervals of less than 100 ms, may completely disappear. The absence of movement in response to the second pulse train is not due simply to fatigue of tectal output neurons or of motor neurons because continuous stimulation of a site for 80 ms elicits a much larger movement that lasts for about 80 ms¹¹.

Although the basis of the refractory effect itself is not understood, we used it to explore the functional organization of the motor control circuitry intermediate to the tectum and spinal cord. When we stimulated two different tectal sites in rapid succession, instead of the second movement disappearing, its

direction was systematically altered. For example, in the case shown in Fig. 1a, the first and second movements when separated by a long interstimulus interval (500 ms) were to the left-up and right-up directions, respectively. As the interval decreased from 500 to 100 ms, however, the direction of the second movement became progressively more rightward (this change in direction is opposite to that expected if the second movement were being partially averaged with the first). Intervals of 100 to 60 ms (the shortest interval used) elicited a second movement whose final position was due right of its starting position. In this case, the upward component of movement was shared by the first and second movements, and it was the upward component that was lost from the second movement, resulting in a purely rightward movement. Thus, the neural circuitry that generates saccadic movements does not become refractory in its entirety. Rather, circuits controlling upward movement can be made refractory independently of circuits controlling rightward movement.

Similarly, circuits that control leftward and upward components of movement were separated by simply reversing the order of stimulation at the same two tectal sites (Fig. 1b). At intervals of less than 100 ms, the direction of the second movement became purely leftward. Analogous results were obtained for paired stimulation sites that evoked movements having a common rightward component (for example, right-up and right-down; Fig. 1c, d): At intervals of less than 100 ms only the downward or upward component remained.

In all of the above cases (29 combinations in 4 owls), the direction of the second movement that remained at short intervals was very close to either a vertical or a horizontal direction. We tested whether these particular directions were dependent on the specific directions of the first and second movements, as

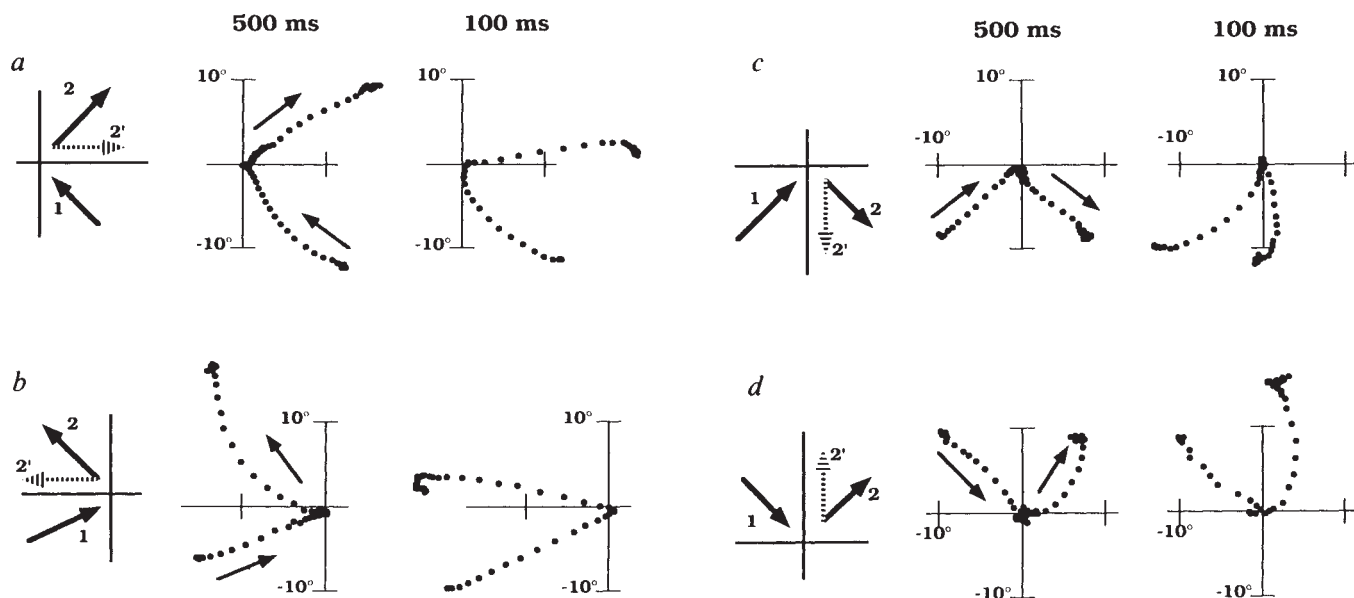


FIG. 1 Examples of head movements elicited by stimulating pairs of tectal sites sequentially at different stimulus intervals. The experimental paradigm was the same as that used by Robinson¹². Each dot represents the position of the head at 5 ms intervals. The axes represent the degree of rotation of the head in the vertical (elevational) and horizontal (azimuthal) directions. Movements were translated according to the position of the head following the first movement; left is left, up is up, etc. *a*, A left-up first movement (1) followed by a right-up second movement at stimulus onset intervals of 500 (2) and 100 (2') ms. At 100 ms, only the rightward component of the second movement persists. *b*, A right-up movement (1) followed by a left-up movement at the same two stimulus onset intervals. At 100 ms, only the leftward component of the second movement persists (2'). *c*, A right-up movement (1) followed by a right-down movement (2 and 2'). *d*, A right-down movement (1) followed by a right-up movement (2 and 2').

METHODS. Owls were prepared for chronic electrode implant under ketamine ($10 \text{ mg kg}^{-1} \text{ h}^{-1}$) and diazepam ($0.6 \text{ mg kg}^{-1} \text{ h}^{-1}$) anaesthesia. Epoxy-

coated tungsten electrodes (tip diameter at base $125 \mu\text{m}$, exposed tip length $250 \mu\text{m}$) plated at the tip with gold and/or platinum (impedance at $1 \text{ kHz} = 10 \text{ k}\Omega$) were lowered into the optic tectum. Visual receptive fields were measured to identify the location of the electrode in the tectum and the shank was cemented to the skull. Four to six electrodes were placed in the two tectal lobes representing sensory field locations (and movement directions) in all spatial quadrants. Several days later the owls were placed in a tube which restricted limb movements but allowed for free head movements. They were positioned at the centre of the induction coils of a 6-ft search coil system (CNC Engineering). Electrical stimulation (constant current, electrically isolated $300 \mu\text{s}$ cathodal pulses, $75\text{--}240 \mu\text{A}$, $40\text{--}60 \text{ ms}$ pulse trains at 300 Hz) was applied through the microelectrodes. The resulting movements were measured from the output of a 2 cm diameter search coil that was attached to the skull. The output of the search coil was recorded, calibrated, and displayed online with a Macintosh II computer, and saved on magnetic disk for subsequent quantitative analysis.

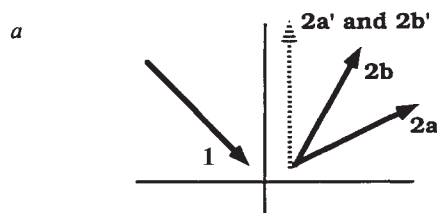
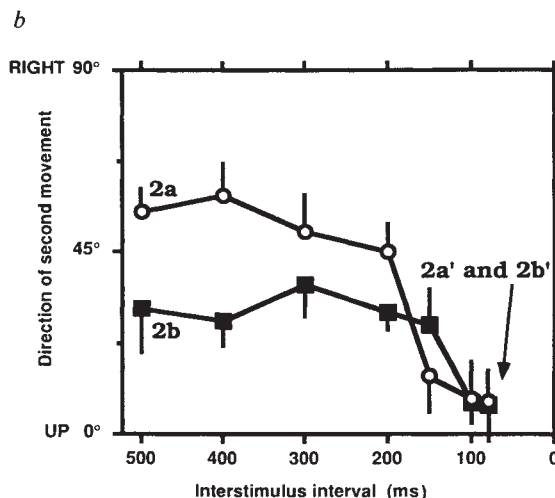


FIG. 2 Two differently directed right-up movements become identically directed upward movements when preceded by a right-down movement. *a*, Effect of the right-down movement (1) on the right-up movements at interstimulus intervals of 500 ms (2a and 2b) and 100 ms (2a' and 2b'). *b*, Mean direction plots (■, ○) of these second movements as a function of interstimulus interval. The first movement was directed to the right-down at 132°. Direction is defined here in a polar coordinate system with 0° representing a purely upward movement and 90° a purely rightward movement. Each symbol represents the mean direction from 5–10 trials; error bars represent standard deviation.



would be predicted by vector addition or by post-tectal circuitry that lacked discrete directional components. To examine this issue, we measured the effect of a single first movement (right-down; movement 1 in Fig. 2a) on two differently directed second movements both into the right-up quadrant of space (movements 2a and 2b in Fig. 2). When either of the two right-up movements was preceded by a movement to the right-down, they changed to become identical, purely upward movements. The reverse strategy yielded the same result: when a single second movement (such as left-down) was tested against two differently directed first movements into a given quadrant of space (such as left-up), the direction of the second movement at short intervals was always purely downward. The same result was observed in each of seven sets of movement combinations. Thus, for movements directed into a given quadrant, the alignment of the refractory movement with one of the orthogonal axes was independent of the direction of the non-refractory movement. These results are incompatible with vector addition, vector averaging, or other models that represent movement direction as a continuous function.

A different phenomenon was observed when two evoked movements were in opposite directions. Under these conditions, the second movement continued to be expressed regardless of the interval. Although intervals of less than 100 ms sometimes caused the first movement to be interrupted by the initiation of

the second, the direction of the second movement was unaltered. Thus, the neural circuits that generate saccades into opposite quadrants of space do not cause each other to become refractory, indicating that they can operate independently of each other.

These results suggest that movement direction is encoded by the coordinated action of a small number of distinct neural circuits (referred to henceforward as saccade generators), each controlling movement in a particular direction. We determined the number of saccade generators and the directions they encode by examining the effect of one movement on another for many combinations of movement directions. Whenever the two movements were into adjacent quadrants of space, the direction of the second movement at short intervals was altered predictably to one of four directions: upward, downward, leftward, or rightward (Fig. 3). These results suggest that the tectal command for movement accesses four independent saccade generators, each responsible for the component of a movement in one of these four orthogonal directions.

The direction of movement evoked by stimulation in one tectal lobe could be altered equally well by stimulation in the same (Fig. 1c, d) or in the opposite tectal lobe (Fig. 1a, b). This implies that the saccade generators are located outside the tectum and that both tectal lobes access the same generators. The existence of a short delay (<20 ms) between tectal stimulation and the onset of movement, combined with the anatomical observations

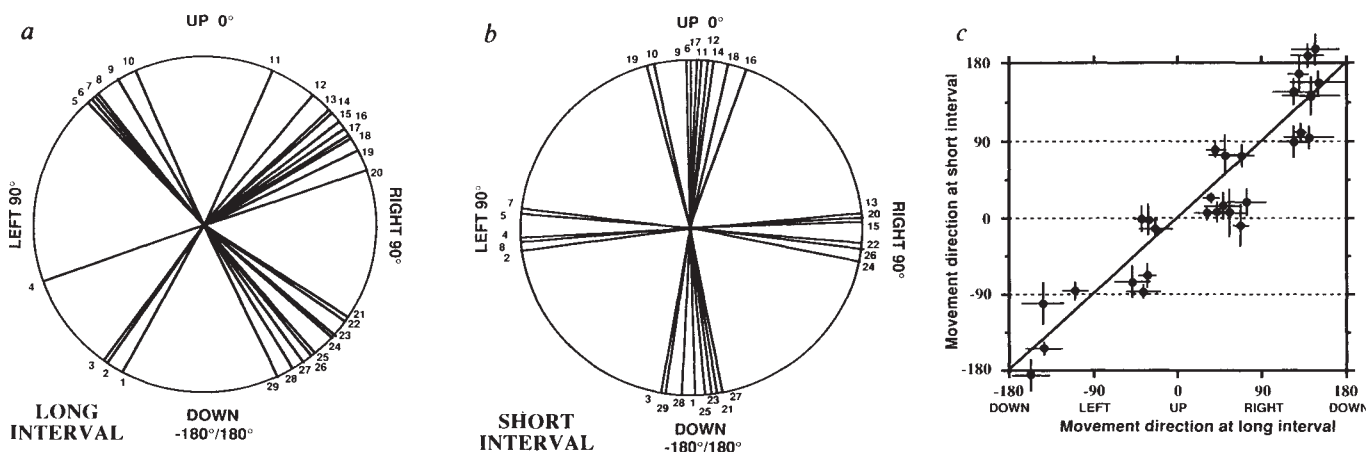


FIG. 3 Summary diagram of the directions of movements elicited from all tectal sites tested when preceded at (a) long (500 ms) and (b) short latency (60–100 ms) by a movement into an adjacent quadrant. Movement direction is represented as a unit vector representing the mean direction of 5–10 movements. Unit vectors are calculated as the direction from starting to the final resting head position: up is up, left is left, etc. Each case is numbered such that, for example, 1 in (a) represents a movement elicited from the

same stimulus site as 1 in (b). *c*, Standard deviations (error bars) associated with the mean movement directions under each condition given in *a* and *b*. The diagonal line represents the hypothesis that the second movement was not affected by the first. The plot also shows that movements in each of the non-orthogonal directions can be altered to either of the two component orthogonal directions.

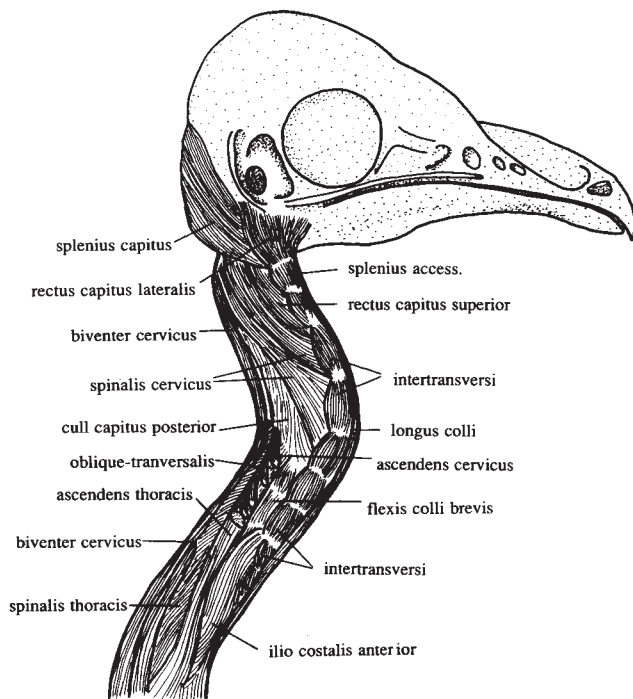


FIG. 4 Diagram of the superficial musculature of the neck of the barn owl (*Tyto alba*). The neck is S-shaped, contains 13 vertebrae and at least 31 distinct muscle pairs, some of which are segmentally arranged whereas others span multiple segments.

on tecto-spinal relay sites in the brainstem tegmentum¹³⁻¹⁵, suggest that the saccade generators are located in brainstem tegmental regions.

The existence of separate circuits for controlling horizontal and vertical components of movement in a complex skeleto-muscular system indicates that such an intermediate representation may be a general principle of sensorimotor organization. In the case of eye movements, the similar orthogonal character of the saccade generators¹² may be related to the essentially orthogonal arrangement of the extraocular rectus muscles. In contrast, head movements depend on the action of many pairs of muscles pulling in a variety of directions, and there is no apparent organization of muscle-pulling directions according to orthogonal planes of movement (Fig. 4). Thus, unlike in eye motor control, the independent coding of orthogonal movement components of the head is markedly different from both sensory and the muscle coordinate frames, implying that it operates as an abstract system intermediate to the sensory and motor processes that control orienting movements.

These observations, together with observations in other systems, suggest that orthogonal coordinate systems may be a general property of intermediate stages in motor control hierarchies. In cats and monkeys, electrical stimulation of several brainstem tegmental regions elicits either primarily horizontal or vertical head movements^{16,17}. Psychophysical experiments on human arm and hand movements suggest that they also may be encoded in terms of their azimuthal and elevational components^{5,18,19}. Lesion studies in the frog have demonstrated that leftward and rightward horizontal components of whole-body orienting movements are controlled by distinct circuitry, and that these circuits are independent of those controlling elevation and distance²⁰⁻²³. This evidence suggests that, despite the differences in the skeletal and muscular organization among head, arm and body movements, there may be for each an intermediate representation of movement direction encoded in orthogonal spatial coordinates.

There may be several advantages to representing movements in orthogonal directional components. The orthogonal coordinate system facilitates the integration of spatial information coming from different sensory modalities, allowing several sensory systems to access a single motor command structure^{18,24-26}. In addition, it facilitates the coordination of movements by different body parts when the goal of a movement can be accomplished in a variety of ways. For example, redirection of gaze can be carried out by movements of the eyes, head or body or by any combination of these three. To locate a desired target accurately, information about the change in gaze contributed by each kind of movement must be combined in a common form before being compared with the change in gaze that was intended. Representing orienting movements in terms of their horizontal and vertical components may provide such a common coordinate frame for feedback regulation of movements by multiple body parts. □

Received 27 November 1989; accepted 26 March 1990.

- Bernstein, N. *The Co-ordination and Regulation of Movements* (Pergamon, Oxford, 1967).
- Hollerbach, J. M. & Flash, T. *Biol. Cybern.* **44**, 67-77 (1982).
- Morasso, P. *Expl. Brain Res.* **42**, 223-237 (1981).
- Bizzi, E., Accornero, N., Chapple, W. & Hogan, N. *J. Neurosci.* **4**, 2738-2744 (1984).
- Soechting, J. F. & Flanders, M. *J. Neurophysiol.* **62**, 595-608 (1989).
- Syka, J. & Radil-Weiss, T. *Brain Res.* **28**, 567-572 (1971).
- Ewert, J.-P. *Brain Behav. Evol.* **3**, 36-56 (1970).
- Schiller, P. J. & Stryker, M. *J. Neurophysiol.* **35**, 915-924 (1972).
- Wurtz, R. H. & Goldberg, M. E. *Science* **171**, 82-84 (1971).
- Sparks, D. L. *Brain Res.* **156**, 1-16 (1978).
- du Lac, S. & Knudsen, E. I. *J. Neurophysiol.* **63**, 131-146 (1989).
- Robinson, D. *Vis. Res.* **12**, 1795-1808 (1972).
- Masino, T. & Knudsen, E. I. *Soc. Neurosci. Abstr.* **14**, 1236 (1988).
- Masino, T. & Grobstein, P. *J. comp. Neurol.* **291**, 103-127 (1990).
- Grantyn, A. & Berthoz, A. in *Control of Head Movement* (eds Peterson B. & Richmond, F.J.) 224-244 (Oxford University Press, 1988).
- Fukushima, K. *Prog. Neurobiol.* **29**, 107-192 (1987).
- Westheimer, G. & Blair, S. M. *Expl. Brain Res.* **24**, 89-95 (1975).
- Soechting, J. F. & Ross, B. *Neuroscience* **13**, 595-604 (1984).
- Lacquaniti, F. *Trends Neurosci.* **12**, 287-291 (1989).
- Kostyk, S. K. & Grobstein, P. *Neuroscience* **21**, 41-55 (1987).
- Grobstein, P. *Brain Behav. Evol.* **31**, 34-48 (1988).
- Masino, T. & Grobstein, P. *Expl. Brain Res.* **75**, 227-244 (1989).
- Masino, T. & Grobstein, P. *Expl. Brain Res.* **75**, 245-264 (1989).
- Knudsen, E. I., du Lac, S. & Esterly, S. A. *Rev. Neurosci.* **10**, 41-65 (1987).
- Simpson, J. I. & Graf, W. *Rev. Oculomotor Res.* **1**, 3-20 (1985).
- Grobstein, P. *Visuomotor Coordination* (eds Ewert, J.-P. & Arbib, M. A.) 537-568 (Plenum, New York, 1989).

ACKNOWLEDGEMENTS. We thank Drs P. Grobstein, W. Newsome, J. Olsen and D. Robinson for their comments on the manuscript. This work was supported by grants from the National Institutes of Health and C. F. Aaron Endowment Fund.

Odour-modulated collective network oscillations of olfactory interneurons in a terrestrial mollusc

Alan Gelperin & David W. Tank

Biophysics Research Department, AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

DETERMINATION of the dynamical structure of neural circuits—the general principles of how neural activity varies with time and manipulates information—is a prerequisite to understanding their computational function¹. Rhythmically active or oscillating neural circuits are particularly interesting dynamical structures, as rhythms and oscillations are a prominent feature of mammalian central nervous system electrophysiology. Coherent oscillations by networks of interneurons are observed in the vertebrate olfactory system^{2,3} and have recently been described in mammalian visual cortex⁴⁻⁶. These interneuronal networks display oscillations in local field potential (LFP) and probability of producing action potentials that are highly correlated between subcircuits sharing

the same stimulus features. Much less is known about the existence and importance of network oscillations in the higher centres of invertebrates⁷. Here we report that a network of olfactory interneurons in the cerebral ganglion of the terrestrial mollusc *Limax maximus* also displays coherent oscillations in LFP which are modified by odour input. This dynamical structure could be central to the odour recognition and odour learning ability of *Limax*^{8,9}.

The network of olfactory interneurons we studied is known as the procerebral (PC) lobe, which contains about 50% of all neurons (10^4 or 10^5 cells) in the entire *Limax* nervous system. The PC lobe receives odour information directly from the superior and inferior noses. Its anatomical features¹⁰⁻¹³ suggest substantial feedback synaptic connections. Extracellular recordings (0–3 Hz bandwidth) from the procerebral lobe of *Limax* (Fig. 1) reveal a periodic potential with a fundamental frequency of 0.69 ± 0.12 Hz ($n = 62$) (Fig. 2a). The LFP oscillation is produced by a mechanism intrinsic to the PC lobe, based on the observations that the oscillation is only recorded by an electrode in the PC lobe and that a PC lobe completely isolated from the rest of the cerebral ganglion still exhibits the oscillation (Fig. 2b). An oscillating LFP is recorded from any point within the PC lobe. Dual microelectrode recordings from within the same PC lobe demonstrate that the oscillations are always phase-locked. Apical and basal recordings within the lobe are in phase but of opposite polarity, whereas the intermediate central region shows a biphasic potential (Fig. 2c). Because the apical and basal regions of the PC lobe show LFP changes of opposite sign, current will flow between them on a cycle by cycle basis, and some of this intra-PC lobe current flow can be sensed by a surface suction electrode (Fig. 2d), directly analogous to surface EEG recordings from mammalian olfactory bulb.

Intracellular recordings from *Limax* PC lobe interneurons demonstrate periodic oscillations of membrane potential (Fig. 2e) with oscillation frequency similar to that of the LFP. During the depolarizing phase of these oscillations, cells can remain silent or can fire one or a short burst of action potentials. Thus the oscillations in *Limax* PC lobe LFP have a correlate in membrane potential oscillations. Extracellular recordings

demonstrate that all spike activity that can be recorded in the PC is phase locked to the LFP oscillation (Fig. 3a, inset). Over 90% of the action potentials occur in the 500 ms preceding the peak of the LFP oscillation (Fig. 3a). Although the ionic currents underlying the LFP and the precise relationship of the LFP to the spike and synaptic activity is at present unclear, the results suggest that the LFP is produced by synaptically-dependent events set up by the coherent collective activity of PC lobe interneurons.

Activation of olfactory afferents by odours under natural conditions will produce input to the PC lobe over several cycles of the LFP oscillation. To determine how the temporal relationship between afferent fibre activation and the LFP oscillation affects the synaptic response produced, the sizes of synaptically dependent evoked potentials were measured for olfactory nerve (ON) shock given at different phases of the LFP oscillation. ON shock produces a two-component response in the neuropil of the PC lobe (Fig. 3b, inset). The first peak is due to the action potentials in the afferent fibres entering the lobe from the ON. The second slower peak is due to synaptic events elicited by the afferent volley⁹. The size of the synaptic response to ON shock is dependent on the time of arrival of the afferent volley relative to the phase of the endogenously oscillating LFP (Fig. 3b). Thus the postsynaptic effects of afferent fibre activation by odour will vary depending upon the phase of the intrinsic PC lobe oscillation at which it arrives. ON shock also transiently alters the PC oscillation frequency and waveform for 5–15 s after the stimulus and resets the rhythm relative to its preshock phase.

The importance of the intrinsic oscillation of the PC lobe in the processing of olfactory information is suggested by experiments using a preparation which retained the olfactory epithelium of the nose connected to the PC lobe via the digitate ganglion and ON¹⁴. The receptor surface was in air while the remainder of the preparation remained immersed in saline. A two-channel puffer was used to apply either moist air or air equilibrated with a $10^{-5}\%$ solution of 2-ethyl-3-methoxypyrazine (potato odour; see ref. 15) to the olfactory surface while recording *en passant* from the olfactory nerve and from the entire PC

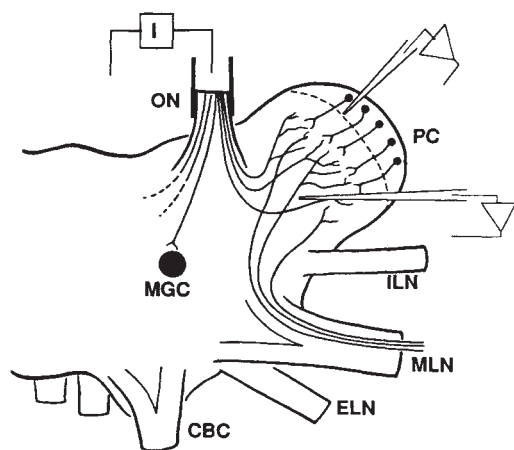


FIG. 1 Schematic diagram of the cerebral ganglion of *Limax maximus*. The olfactory nerve (ON) is in a suction electrode for stimulation. Extracellular micropipette electrodes are shown recording from apical (cell body) region and basal neuropil in the procerebral (PC) lobe. Specimens of *Limax maximus* ($n = 74$) were anaesthetized by chilling on ice for 30 min before the brain was removed into cold (4–6 °C) saline for isolation of the cerebral ganglia. Isolated PC lobe preparations were made by cutting across the base of the lobe from the insertion of the ON to the insertion of the internal lip nerve (ILN). Nose-brain preparations retained the olfactory neuroepithelium and digitate ganglion from the tip of the superior tentacle connected to the ON. The PC lobe also receives input from the inferior nose via the medial lip nerve (MLN). CBC-cerebrobuccal connective, ELM-external lip nerve, MGC-metacerebral giant cell.

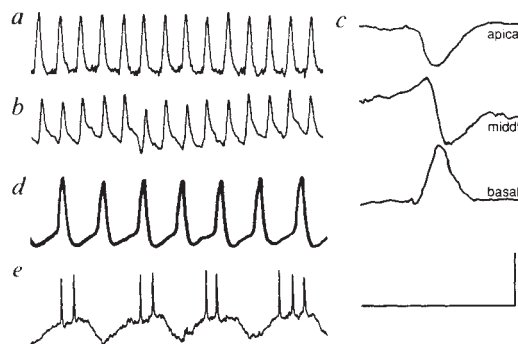


FIG. 2 Recordings of spontaneous local field potential (LFP) oscillations from a PC lobe. a, PC lobe attached to the cerebral ganglion as in Fig. 1. Recording site located in the neuropil. b, Same recording site in same lobe as (a) but after complete isolation from the cerebral ganglion. c, Simultaneous recordings from apical, central and basal recording sites in a PC lobe. d, Coherent PC lobe oscillation recorded from an isolated lobe lodged in a suction electrode. e, Intracellular recording from a neuron in the PC lobe. Calibration bars are 10 s, 1 pA (a, b); 700 ms, 3 pA (c); 10 s, 200 μ V (d); 2.5 sec, 50 mV (e). Extracellular recordings were made either with a List model EPC 5 or a Getting model 5A amplifier and were band-pass filtered between 0–3 Hz using a PAR model 113. Data were recorded on a Gould model 220 chartwriter or digitized using a MacPacq system and stored on an Apple MacII computer for later analysis. Intracellular recordings were made using beveled glass microelectrodes of 80–120 M Ω resistance and a NeuroData IR-283 amplifier.

lobe with a suction electrode. A 10 s puff of potato odour activated many small amplitude units and a few large amplitude units in the ON and dramatically altered the waveform and frequency of the PC lobe oscillation (Fig. 4a). A puff of moist air at the same velocity elicited much less activity in both the ON and in the PC lobe (Fig. 4b). The response during odour application is predominantly a transient reduction in amplitude ($15.4 \pm 2.3\%$, $n = 14$) (mean $\pm$ s.e.m.) and reduction in frequency ($9.2 \pm 2.8\%$, $n = 15$) of the endogenous oscillation. This pattern of response was seen in 26 of 29 trials with 5 nose-brain

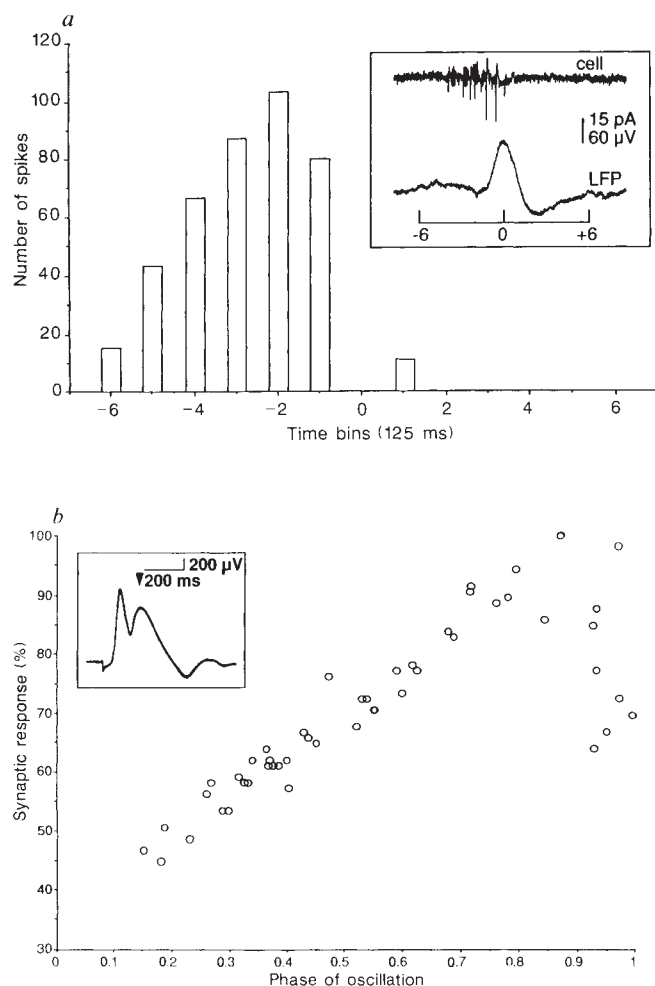


FIG. 3 Spontaneous oscillation phase data. *a*, Extracellular cell-attached patch electrode recordings ($n = 45$) showing spike activity phase-locked to the spontaneously oscillating local field potential (LFP). Inset: Typical record of multiunit extracellular recording and simultaneously recorded local field potential. Calibration bar is 15 pA for cells, 60 μ V for LFP. Histogram shows number of action potentials occurring in 125 ms time bins with zero time taken at the peak of the LFP. *b*, Amplitude of the synaptic current evoked in the PC lobe by supramaximal shock of the ON depends on the phase of the spontaneous oscillation at which it is evoked. Inset: the two-phase response to a single ON shock. The second peak of the response is the calcium-sensitive synaptically-dependent event. Calibration bars indicate 150 ms, 300 μ V. The graph shows that the evoked synaptic response amplitude depends on spontaneous oscillation phase. Data samples (5 s) were obtained every 100 s, with ON shock occurring 4 s into the sample. The time of the last complete oscillation cycle before the shock (T_1) was measured, as was the time from the last oscillation peak before the shock to the peak of the synaptic response (T_2). The phase of occurrence of the synaptic response was taken as (T_2/T_1) . The evoked synaptically-dependent event is minimal just before and after the occurrence of the peak of the spontaneous oscillation and then increases smoothly in amplitude over the middle 80% of the inter-peak interval.

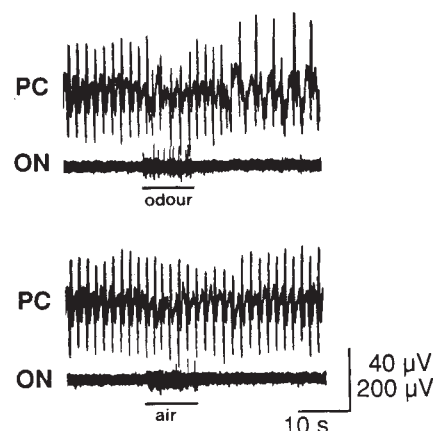


FIG. 4 Response to potato odour recorded using a nose-brain preparation. Suction electrodes recorded *en passant* from the ON and from the surface of the PC lobe. *a*, A 10-s puff of air saturated with the vapour of $10^{-5}\%$ (v/v) of 2-ethyl-3-methoxypyrazine applied to the receptor epithelium of the nose elicited multiunit activity in the ON and dramatically altered the waveform and frequency of the PC lobe oscillation. *b*, A 10-s puff of moist air at the same velocity produced much less activity in the ON and a smaller alteration in the PC lobe LFP oscillation. The change in oscillation frequency following odour offset shown in (*a*) was not observed in every preparation.

preparations. These results indicate that activation of one of the normal sensory input pathways to the PC lobe using a behaviourally relevant stimulus at a concentration known to be attractive to the intact animal¹⁵ produces, among other effects, transient changes in the endogenously active local circuit responsible for the oscillating LFP.

The existence in *Limax* of the oscillating LFP and its responsivity to ON shock and odour-elicited input from the olfactory epithelium establishes yet another functional parallel between the molluscan and mammalian olfactory systems. A striking feature of mammalian olfactory bulb is the endogenously oscillating 40 Hz LFP that is enhanced by inspiration of odours¹⁶⁻²⁰. The amplitude of the LFP oscillation varies over the bulbar surface. The spatial pattern of oscillation amplitudes could represent an odour and can be altered by conditioning²¹⁻²³. A theoretical model has explored how the LFP oscillation by the olfactory bulb may contribute to the signal processing functions in the detection of weak odours²⁴. Different models have explored how the olfactory bulb²⁵ or olfactory (piriform) cortex²⁶⁻²⁸ could implement associative (content-addressable) memories. The possibility that endogenous oscillations in the second-order olfactory circuit of *Limax* allow detection of weaker odour inputs than would otherwise be the case has been supported by recent studies of LIMAX, a neural network model of the *Limax* olfactory processing and odour learning circuit²⁹. With the identification of an output pathway from the PC lobe to the pedal ganglia¹² it may be possible to test this idea directly. The rhythmic synchrony of PC lobe cell activity coupled throughout the entire lobe must have a critical role in the odour processing function of this structure. The striking functional analogy between such phylogenetically disparate systems indicates that the computational use of the local circuit oscillation is fundamental to the tasks of these types of olfactory processing networks. Other olfactory systems, for example the central olfactory pathways of insects^{30,31}, may display similar dynamics.

The oscillating LFP in *Limax* PC lobe is modulated in both waveform and frequency by dopamine and serotonin³² (A.G., L. Rhines, J. Flores, and D.W.T., unpublished data), both of which are present in fibres entering the lobe from sources extrinsic to the lobe^{33,34}. Both serotonin, previously implicated in mechanisms of molluscan synaptic plasticity³⁵⁻³⁹, and dopamine alter the phosphorylation state and rate of synthesis

of specific proteins in neurons of the PC lobe^{40,41}. Therefore, as in mammalian olfactory bulb⁴²⁻⁴⁴, input pathways exist that could modify PC lobe circuitry during aversive odour conditioning. □

Received 20 November, 1989; accepted 13 March, 1990.

1. Tank, D. W. *Semin. Neurosci.* **1**, 67-79 (1989).
2. Freeman, W. J. *Mass Action in the Nervous System* (Academic, New York, 1975).
3. Shepherd, G. M. in *Taste, Olfaction and the Central Nervous System* (ed. Pfaff, D. W.) 307-332 (Rockefeller University Press, New York, 1985).
4. Gray, C. M. & Singer, W. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1698-1702 (1989).
5. Gray, C. M., König, P., Engel, A. K. & Singer, W. *Nature* **338**, 334-337 (1989).
6. Eckhorn, R. et al. *Biol. Cybern.* **60**, 121-130 (1988).
7. Bullock, T. H. & Basar, E. *Brain Res. Rev.* **13**, 57-75 (1988).
8. Gelperin, A., Tank, D. W. & Tesaro, G. in *Neural Models of Plasticity* (eds Byrne, J. H. & Berry, W. O.) 133-159 (Academic, New York, 1989).
9. Sahley, C. L. in *Connectionist Modeling and Brain Function: The Developing Interface* (eds Hanson, S. & Olson, C.) 36-73 (MIT, Cambridge, 1990).
10. McCarragher, G. & Chase, R. *J. Neurobiol.* **16**, 69-74 (1985).
11. Zs.-Nagy, I. & Sakharov, D. A. *Tissue & Cell* **2**, 399-411 (1970).
12. Chase, R. & Tolloczko, B. *J. comp. Neurol.* **283**, 143-152 (1989).
13. Veratti, E. *Mem. Real. Inst. Lomb. Sci. Lett.* **18**, 163-179 (1900).
14. Egan, M. E. & Gelperin, A. *J. mollusc. Stud.* **47**, 80-88 (1981).
15. Hopfield, J. F. & Gelperin, A. *Behav. Neurosci.* **103**, 329-333 (1989).
16. Freeman, W. J. *Biol. Cybern.* **35**, 221-234 (1979).
17. Adrian, E. D. *J. Physiol.* **100**, 459-473 (1942).
18. Adrian, E. D. *Electroenceph. clin. Neurophysiol.* **2**, 377-388 (1950).
19. Adrian, E. D. *Br. med. Bull.* **6**, 330-333 (1950).
20. Freeman, W. J. & Skarda, C. A. *Brain Res. Rev.* **10**, 147-175 (1985).

21. Gray, C. M., Freeman, W. J. & Skinner, J. E. *Behav. Neurosci.* **100**, 585-596 (1986).
22. Freeman, W. J. & Viana Di Prisco, G. *Behav. Neurosci.* **100**, 753-763 (1986).
23. Baird, B. *Physica* **22D**, 150-175 (1986).
24. Li, Z. & Hopfield, J. J. *Biol. Cybern.* **61**, 379-392 (1989).
25. Freeman, W. J., Yao, Y. & Burke, B. *Neur. Net.* **1**, 277-288 (1988).
26. Wilson, M. A. & Bower, J. M. in *Neural Information Processing Systems* (ed. Anderson, D. Z.) 114-126 (American Institute of Physics, New York, 1988).
27. Lynch, G., Granger, R., Baudry, M. & Larson, J. in *Neural Connections, Mental Computation* (eds Nadel, L., Cooper, L., Culicover, P. & Harnish, R. M.) 247-289 (MIT Press, Cambridge, 1988).
28. Haberly, L. B. *Chem. Senses* **10**, 219-238 (1985).
29. Hopfield, J. J. in *Computer Simulation in Brain Science* (ed. Cotterill, R. M. J.) 405-415 (Cambridge University Press, New York, 1988).
30. Homberg, U., Christensen, T. A. & Hildebrand, J. G. A. *Rev. Ent.* **34**, 477-501 (1989).
31. Menzel, R. in *Perspectives in Neural Systems and Behavior* (eds Carew, T. & Kelley, D.) 249-266 (A. R. Liss, New York, 1989).
32. Rhines, L. thesis, Princeton Univ. (1989).
33. Osborne, N. N. & Cottrell, G. A. *Z. Zellforsch. mitrosk. Anat.* **112**, 15-30 (1971).
34. Yamane, T., Gelperin, A. & Delaney, K. *Soc. Neurosci. Abstr.* **12**, 862 (1986).
35. Kandel, E. R. & Schwartz, J. H. *Science* **218**, 433-443 (1982).
36. Sweatt, J. D. & Kandel, E. R. *Nature* **339**, 51-54 (1989).
37. Crow, T. *Trends Neurosci.* **11**, 136-142 (1988).
38. Eskin, A., Garcia, K. S. & Byrne, J. H. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2458-2462 (1989).
39. Grover, L. M., Farley, J. & Auerbach, S. B. *Brain Res. Bull.* **22**, 363-372 (1989).
40. Yamane, T. & Gelperin, A. *Cell. mol. Neurobiol.* **7**, 291-301 (1987).
41. Yamane, T., Oestreicher, A. B. & Gelperin, A. *Cell. mol. Neurobiol.* **9**, 447-459 (1989).
42. Macrides, F., Davis, B. J., Young, W. M., Nadi, N. S. & Margolis, F. L. *J. comp. Neurol.* **203**, 495-514 (1981).
43. Shipley, M. T., Halloran, F. J. & de la Torre, J. *Brain Res.* **329**, 294-299 (1985).
44. McLean, J. H. & Shipley, M. T. *J. Neurosci.* **7**, 3016-3028 (1987).

ACKNOWLEDGEMENTS. We thank Jorge Flores for help with data collection and John Hopfield, Alan Kay, Wade Regehr and Kerry Delaney for comments on the manuscript.

Ciliary neurotrophic factor prevents the degeneration of motor neurons after axotomy

Michael Sendtner*, Georg W. Kreutzberg† & Hans Thoenen

Departments of Neurochemistry and † Neuromorphology, Max-Planck-Institute for Psychiatry, D-8033 Planegg-Martinsried, FRG

THE period of natural cell death in the development of rodent motor neurons is followed by a period of sensitivity to axonal injury¹⁻³. In the rat this early postnatal period of vulnerability coincides with that of very low ciliary neurotrophic factor (CNTF) levels in the sciatic nerve before CNTF increases to the high, adult levels⁴. The developmental time course of CNTF expression, its regional tissue distribution and its cytosolic localization (as suggested by its primary structure)^{4,5} favour a role for CNTF as a lesion factor rather than a target-derived neurotrophic molecule like nerve growth factor. Nevertheless CNTF exhibits neurotrophic activity *in vitro* on different populations of embryonic neurons⁶. To determine whether the vulnerability of motor neurons to axotomy in the early postnatal phase is due to insufficient availability of CNTF, we transected the axons of newborn rat motor neurons and demonstrated that local application of CNTF prevents the degeneration of the corresponding cell bodies.

We have chosen the rat facial nerve as an experimental system as it exclusively contains motor axons distal to the stylomastoid foramen. The corresponding cell bodies comprise a homogeneous, well-delineated nucleus in the ventral-lateral region of the brain stem (Fig. 1). Moreover, the retrograde and degenerative changes occurring after facial-nerve lesion have been investigated in great detail⁷. The facial nerves of newborn rats were unilaterally transected on the first postnatal day (see Table 1) and morphology of the facial nuclei was analysed 1 week after lesion. More than 80% of the cell bodies on the side having the lesion were completely lost after 1 week (Table 1). The remaining motor neurons showed severe signs of degeneration⁸, that is, the nuclei were located eccentrically and the Nissl bodies showed the characteristic dispersion that occurs in adult motor neurons after axotomy (Fig. 1). In adult animals astrocytes

TABLE 1 Motor neuron survival after lesion of the facial nerve in newborn rats

	Number of surviving neurons ± s.e.m.	
	Side with lesion	Side without lesion
Untreated animals (n=1)	685	2,985
Animals treated with BSA gel foam (n=3)	620 ± 98*	3,271 ± 61
Animals treated with gel foam containing 5 µg of CNTF (n=4)	2,503 ± 487*	3,281 ± 112

Newborn Wistar rats were anaesthetized by hypothermia and the right facial nerves cut ~1 mm distal to the stylomastoid foramen. Gel foam (Spongostan, gift of K. Unsicker, Marburg) was cut into small pieces (3 mm long) and soaked in PBS buffer (20 µl) containing 5 µg purified adult rat sciatic nerve CNTF^{4,13}. The soaked foam pieces were inserted at the sites of lesion. Control animals were treated with gel foam soaked in PBS buffer containing 5 µg BSA (Fraction V, Sigma) (20 µl). The skin was sutured with silk (Ethicon 3-0) and the animals returned to their mother. Operated rats were inspected daily, lesion of the right facial nerve was detected in the animals by lack of movement of the right whisker and the right corner of the mouth. On postnatal day 7 the animals were anesthetized with ether and killed by transcardial perfusion with 4% formaldehyde. The brainstem was dissected, postfixed (1 h), rinsed in water, dehydrated with increasing concentrations of ethanol (70-100%) and embedded in paraffin. Serial sections 7 µm were made from the whole brainstem. After Nissl staining the motoneurons of both facial nuclei were clearly detectable. Only cells containing a clearly-visible nucleolus were counted, in every fifth section, as described previously¹⁴.

* $P < 0.05$, t-test

react characteristically to peripheral axonal lesion of motor neurons by increasing their glial fibrillary acidic protein (GFAP) production^{9,10}. In agreement with this there was enhanced GFAP staining in the astrocytes surrounding the motor neurons on the side having the lesion (data not shown). This is noteworthy because, at the end of the first postnatal week there is normally little GFAP reactivity in the brain stem; it only becomes weakly apparent at later developmental stages¹¹.

The application of CNTF to the proximal stump of the nerve in which the lesion had been made (5 µg of highly purified CNTF⁴ corresponding to 50,000 TU of biological activity) resulted in the survival of most of the motor neurons up to 1 week after lesion (Table 1). Morphological degenerative changes in

* To whom correspondence should be addressed.

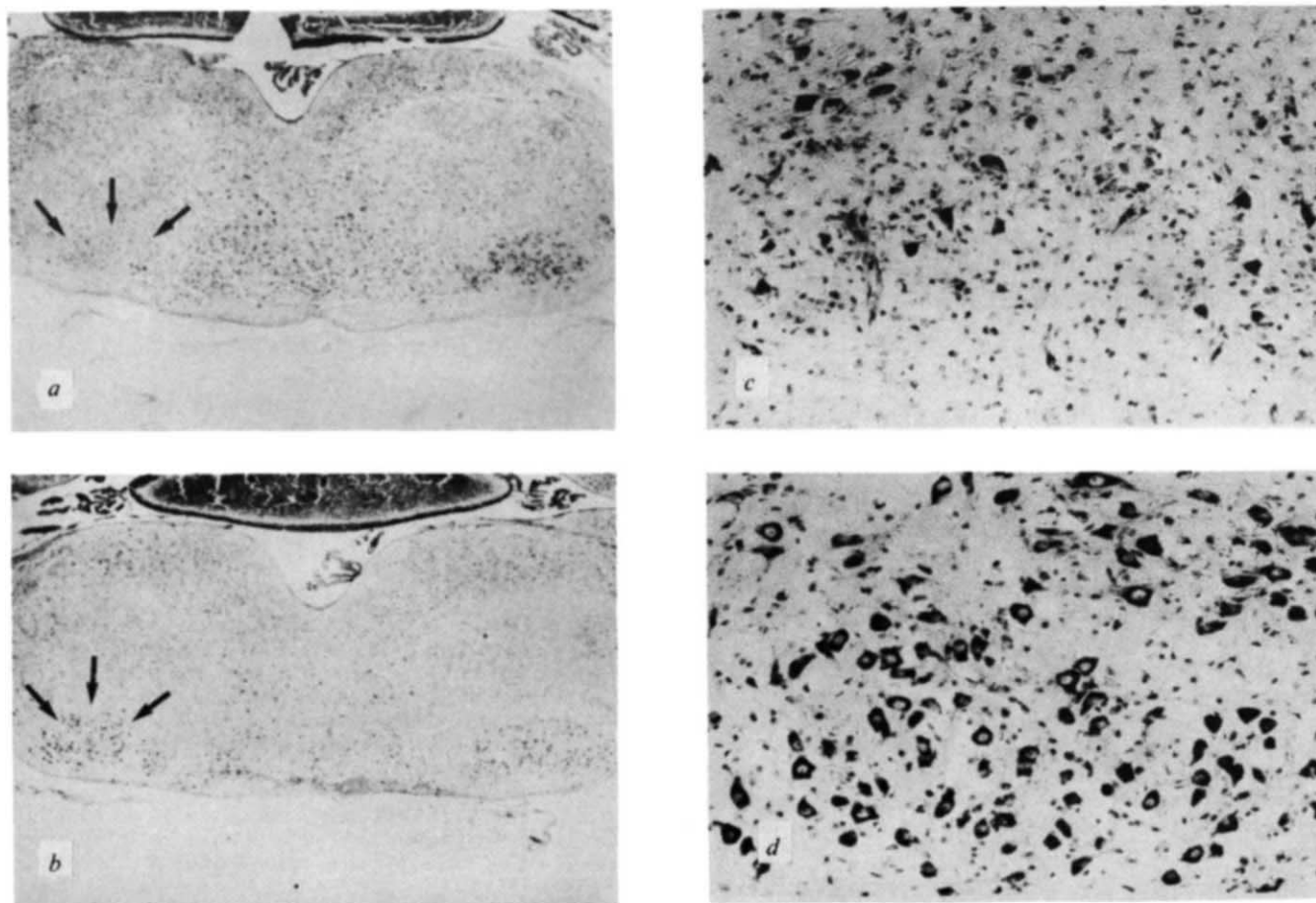


FIG. 1 *a, b*, Sections of the brainstem of 7-day-old rats: The facial nucleus of the side without a lesion is located on the righthand side. The facial nucleus containing the cell bodies of the neurons in which lesions have been made are located on the lefthand side (arrows). *a*, Control animal after insertion of gel foam containing 5 μ g of BSA. *b*, CNTF-treated animal. (Magnification $\times 25$). *c, d*, Cell bodies of the facial nucleus seven days after

lesion. *c*, Control animal treated with BSA. *d*, CNTF-treated animal. (Magnification $\times 250$). The nuclei of the motor neurons in which a lesion has been made in BSA-treated animals (*c*) are deformed and located peripherally (a typical sign of chromatolysis), whereas in CNTF-treated animals most nuclei appear round in shape and are located in the centre of the cell (*d*).

facial neurons in which a lesion had been made were also markedly reduced compared to those observed in the small number of surviving neurons where the lesion had been made and the animals BSA-treated (Fig. 1*c, d*). Despite the survival of most motor neurons in which a lesion had been made, there was still an enhanced GFAP reactivity in the facial nucleus of CNTF-treated nerves (data not shown), showing that CNTF can prevent the degenerative changes in motor neurons (although the reactive changes in astrocytes were still clearly visible). Immunohistochemical procedures¹² did not show whether there was a quantitative difference in the GFAP staining between the CNTF- and BSA-treated animals.

These experiments show that the local application of purified CNTF prevents the lesion-induced death of motor neurons in the rat facial brain stem nucleus. This observation is consistent with our hypothesis that in adult animals the CNTF—present in large quantities in non-neuronal cells of the peripheral nerves—is released after lesion in quantities sufficient to prevent the degeneration of motor neurons in which a lesion has been made. In contrast, in the early postnatal period the quantities of CNTF available are normally insufficient to prevent this, unless exogenous CNTF is applied locally. Although our experiments do not indicate whether the protective effects of locally applied CNTF are due to direct or indirect action on motor neurons, CNTF does support the survival of motor neurons during the naturally occurring period of cell death (Arakawa

and M.S., unpublished observation). Thus, CNTF may act as a general neuronal-survival factor unrestricted in its time of action.

The observation that motor neurons' death can be prevented after lesion in the early postnatal period could introduce novel therapeutic approaches to the treatment of motor neuron degeneration. For example, it will be of interest to investigate whether CNTF also protects against toxic or genetic degenerative changes in motor neurons, and possibly also in those motor neuron diseases of unknown aetiology. □

Received 18 January; accepted 28 March 1990.

1. Sheard, P., McCaig, C. D. & Harris, A. J. *Devl Biol.* **102**, 21–31 (1984).
2. Schmalbruch, H. J. *comp. Neurol.* **224**, 252–258 (1984).
3. Gudden, B. *Arch. Psychiat. Nervenkrankh.* **2**, 693–723 (1870).
4. Stöckli, K. A. *et al. Nature* **342**, 920–923 (1989).
5. Lin, L.-F. H. *et al. Science* **246**, 1023–1025 (1989).
6. Barbin, G., Manthorpe, M. & Varon, S. *J. Neurochem.* **43**, 1468–1478 (1984).
7. Kreutzberg, G. W. in *The Facial Nerve* (ed. May, M.) 75–83 (Thieme, New York, 1986).
8. Nissl, F. *Allg. Z. Psychiat.* **48**, 197–198 (1892).
9. Graeber, M. B. & Kreutzberg, G. W. *J. Neurocytol.* **15**, 363–373 (1986).
10. Tetzlaff, W., Graeber, M., Bisby, M. A. & Kreutzberg, G. W. *Glia* **1**, 90–95 (1988).
11. Bignami, A. & Dahl, D. *Nature* **252**, 55–56 (1974).
12. Springall, D. R., Hacker, G. W., Grimelius, L. & Polak, J. M. *Histochemistry* **81**, 603–608 (1984).
13. Ernsberger, U., Sendtner, M. & Rohrer, H. *Neuron* **2**, 1275–1284 (1989).
14. Oppenheim, R. W. *J. comp. Neurol.* **246**, 281–286 (1986).

ACKNOWLEDGEMENTS. We thank David Edgar for critical reading of the manuscript, Waltraud Komp and Dietmute Bueringer for technical assistance, and Elfi Grossmann for typing the manuscript.

The murine mutation osteopetrosis is in the coding region of the macrophage colony stimulating factor gene

Hisahiro Yoshida*†, Shin-Ichi Hayashi*‡, Takahiro Kunisada*, Minetaro Ogawa*, Satomi Nishikawa*, Hitoshi Okamura†, Tetsuo Sudo§, Leonard D. Shultz|| & Shin-Ichi Nishikawa*

* Department of Pathology, Institute for Medical Immunology, and

† Department of Obstetrics and Gynecology, Kumamoto University Medical School, 2-2-1 Honjo, Kumamoto 860, Japan

§ Biomaterial Research Institute Co. Ltd, Yokohama, Kanagawa 244, Japan

|| Jackson Laboratory, Bar Harbor, Maine 04609, USA

MICE homozygous for the recessive mutation osteopetrosis (*op*) on chromosome 3 have a restricted capacity for bone remodelling, and are severely deficient in mature macrophages and osteoclasts¹⁻³. Both cell populations originate from a common haemopoietic progenitor. As *op/op* mice are not cured by transplants of normal bone marrow cells⁴, the defects in *op/op* mice may be associated with an abnormal haematopoietic microenvironment rather than with an intrinsic defect in haematopoietic progenitors. To investigate the molecular and biochemical basis of the defects caused by the *op* mutation, we established primary fibroblast cell lines from *op/op* mice and tested the ability of these cell lines to support the proliferation of macrophage progenitors. We show that *op/op* fibroblasts are defective in production of functional macrophage colony-stimulating factor (M-CSF), although its messenger RNA (*Csfm* mRNA) is present at normal levels. This defect in M-CSF production and the recent mapping of the *Csfm* structural gene near *op* on chromosome 3 (refs 5, 6) suggest that *op* is a mutation within the *Csfm* gene itself. We have sequenced *Csfm* complementary DNA prepared from *op/op* fibroblasts and found a single base pair insertion in the coding region of the *Csfm* gene that generates a stop codon 21 base pairs downstream. Thus, the *op* mutation is within the *Csfm* coding region and we conclude that the pathological changes in this mutant result from the absence of M-CSF.

Previous reports show that macrophage numbers are markedly reduced in various tissues from *op/op* mice². In order to investigate the molecular basis of the macrophage defect in *op/op* mice *in vitro*, lung fibroblast lines were established from B6C3-*op/op* and +/? normal littermate mice. These fibroblast lines were tested for their ability to produce factors that support the proliferation of normal macrophage precursors. BALB/c bone marrow cells (passed twice through Sephadex G-10 columns) were co-cultured with the fibroblast lines for two weeks. Massive proliferation of haemopoietic cells occurred with fibroblasts from normal littermate mice (Fig. 1a), and the majority of recovered cells (98.5 ± 1.2%) were macrophages (Fig. 1b). In contrast, proliferation of bone marrow cells was severely reduced with fibroblasts from *op/op* mice and the majority of recovered cells were granulocytes, although a small percentage of macrophages were also detected (5.8 ± 0.9%) (Fig. 1c). This result shows that the macrophage microenvironment defect of *op/op* mice is reproduced *in vitro* using the lung fibroblast culture system.

To determine whether the failure of *op/op* fibroblast cultures to support normal macrophage development is due to a defect in the production of growth factor(s) which induce the prolifer-

ation of macrophage precursors, colony-stimulating activity in the culture supernatant of *op/op* fibroblasts was assessed. It induced a fifth of the number of colonies induced by culture supernatant from normal fibroblasts. Moreover, anti-GM-CSF

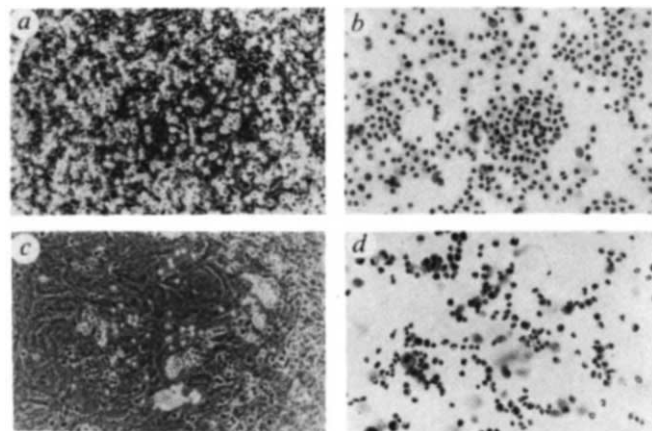


FIG. 1 Low efficiency of *op/op* fibroblasts in supporting the proliferation of macrophage precursors. BALB/c bone marrow cells (depleted of stromal cells by passage through Sephadex G-10 columns) were cultured either on *op/op* or +/?-littermate fibroblasts for 2 weeks. a, Phase contrast view of the culture on littermate fibroblasts and c, on *op/op* fibroblasts. May-Grunwald-Giemsa staining of the cells present in the culture of b, littermate fibroblasts and d, *op/op* fibroblasts.

METHODS. Preparation of primary fibroblast cell lines: Lung tissues were minced and incubated with 500 U ml⁻¹ collagenase (Sigma) in RPMI 1640 medium (GIBCO) containing 5% calf serum (Hyclone), 5 × 10⁻⁵ M 2-mercaptoethanol (Sigma), 50 µg ml⁻¹ streptomycin and 50 U ml⁻¹ penicillin (Meiji Seika, Tokyo) in 5% CO₂ at 37 °C for 12 h. After washing three times with the same medium, the total lung cells from individual mice were seeded in 100-mm tissue culture dishes (Corning). The cultured cells were grown to confluency, collected by trypsin-EDTA treatment, and subcultured into six dishes. This protocol was repeated over 10 times, and the cells obtained were used as fibroblast cell lines in this study. We prepared fibroblast lines from three pairs of *op/op* and littermate +/? mice at 3-, 4-, and 6-weeks of age. Preparation of bone marrow cells and co-culture with fibroblasts: To remove mature macrophages and adherent stromal cells, BALB/c bone marrow cells were passed through a Sephadex G-10 column twice (100 × 10⁶ cells per 20 ml, Pharmacia) and 10⁶ of the eluted cells were seeded on confluent fibroblast layers in T25 flasks (Corning). The medium (5% calf serum, 5 × 10⁻⁵ 2-mercaptoethanol, RPMI 1640 (6 ml)) was changed every three days.

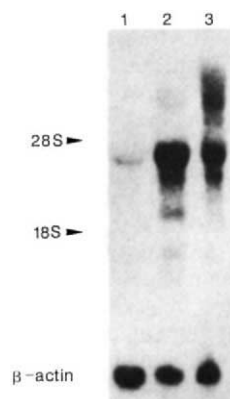


FIG. 2 Expression of M-CSF transcripts in *op/op* fibroblasts. Total RNA was prepared from *op/op* (lane 3), +/?-littermate (lane 2), fibroblasts or EL4 thymoma cell lines (lane 1). 20 µg of each RNA was separated by PAGE and blotted onto nitrocellulose membrane (Hybond-C; NEN). A full-length *Csfm* cDNA clone (CDM8MCSFCL7) was used as a hybridization probe. The murine β -actin probe used as a control indicates equal loading.

† To whom correspondence should be addressed.

antisera abolished all the colony-stimulating activity of *op/op* fibroblasts. Thus, GM-CSF was the only detectable colony-stimulating activity secreted by the *op/op* fibroblasts (Table 1). In contrast, 80% of colonies induced by the culture supernatant of normal littermate fibroblasts remained after anti-GM-CSF antiserum treatment. The macrophage-inducing factors expressed in normal fibroblasts are GM-CSF and M-CSF (refs 7-9). Our results show that M-CSF production is defective in *op/op* mice.

To test whether *op/op* fibroblasts are defective in *Csfm* mRNA production, mRNA was prepared from *op/op* or littermate fibroblasts and expression of the *Csfm* gene measured by northern blot analysis. Despite the absence of functional M-CSF activity in *op/op* fibroblasts, the level of *Csfm* mRNA expression was normal, indicating that it is the protein that is aberrant in *op/op* mice (Fig. 2).

To characterize the mutation in the *Csfm* gene of *op/op* mice we prepared cDNA from *op/op* fibroblast lines and amplified the gene using the polymerase chain reaction¹⁰ (PCR), before cloning into Bluescript (pBS) M13⁺ vectors. As M-CSF activity is present in the N-terminal third of the *Csfm* gene¹¹⁻¹⁴ only cDNA corresponding to this region was sequenced. Four independent clones had a T (thymidine) insertion 262 base pairs (bp) downstream from the ATG initiation codon by comparison with *Csfm* cDNA cloned from mouse L-cells¹². No such insertion was detected in the *Csfm* cDNA from a normal stromal cell line, ST2, which secretes functional M-CSF (refs 15, 16; Fig. 3). Because this insertion generates a TGA stop codon 21 bp

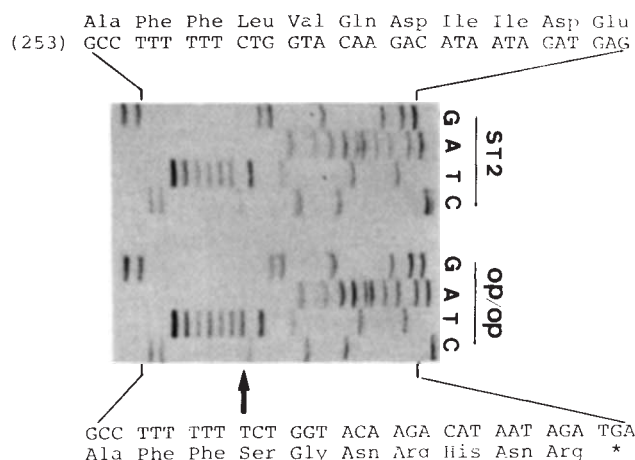


FIG. 3 One-base pair insertion in the coding region of the *Csfm* gene of *op/op* mutant. The coding region of the *Csfm* gene expressed in *op/op* fibroblasts was cloned using PCR and the nucleotide sequence was determined. Nucleotide sequence of *Csfm* cDNA clone at the site of mutation; *op/op* (clone pBSop-1) and ST2 (pBSST2). The same insertion was detected in three additional independent clones from *op/op* fibroblasts, whereas no such insertion was detected in four independent clones from normal fibroblasts.

METHODS. cDNA was prepared from 10 µg total RNA using a cDNA synthesis kit (BRL) with oligo(dT) primer and used as the template for PCR amplification of the *Csfm* coding region¹⁰. Oligonucleotide primers for PCR amplification were CCGGGATCCAGCTGCCCGT (5' primer) and ATAGAAATCTTTCTTACTAGGAGTTC (3' primer), synthesized by an automated DNA synthesizer (ABI). The conditions used were: 2.5 U *Taq* DNA polymerase (BRL), 100 pmol of primers in 50 µl 2 mM MgCl₂, 50 mM KCl, 25 mM TAPS buffer (pH 9.3), 1 mM DTT and 30 cycles in a Program Temp Control System PC-500 (ASTECH, Kitakyushu, Japan). Amplified genes were digested with *Hind*III and *Eco*RI restriction enzymes (NEB), purified by agarose gel electrophoresis, cloned into pBSM13⁺ (Stratagene) and sequenced. Primers used for sequence reactions were GAACTCTCCAATAACCTG, GAACAGCTGGATGATCCTG and CAGGATCATCCAGCTGTTC (which are the sequences in the coding region of *Csfm* gene) and T3 and T7 primers (Stratagene).

TABLE 1 Absence of CSF-production other than GM-CSF from *op/op* fibroblasts

Factors		No. of colonies per 10 ⁵ bone marrow cells			
		Medium	NRS	Anti-IL-3	Anti-GM-CSF
Medium		0	0	0	0
IL-3	(100 U)	199 ± 1	213 ± 52	0	213 ± 1
GM-CSF	(100 U)	149 ± 15	146 ± 24	170 ± 15	0
LCCM	(10%)	112 ± 12	89 ± 3	98 ± 5	76 ± 20
+/? SN	(10%)	210 ± 13	214 ± 12	211 ± 25	174 ± 19
	(2%)	164 ± 26	164 ± 10	160 ± 18	131 ± 21
<i>op/op</i> SN	(10%)	44 ± 12	42 ± 8	41 ± 1	3 ± 3
	(2%)	3 ± 1	1 ± 1	2 ± 1	0

Colony formation assays in soft agar (DIFCO) were carried out as described¹⁹. Positive controls were: recombinant mouse IL-3, GM-CSF and L-cell conditioned medium (LCCM) as a source of M-CSF. Tissue culture supernatant was prepared from primary lung-fibroblast cell lines derived from +/? littermates or *op/op* mice by incubating confluent fibroblast cultures for five days in fresh medium. To identify the factors secreted into the culture supernatant, we used rabbit antisera against murine GM-CSF or IL-3. We used 1% dilutions of the antibodies, which can saturate ~4 U ml⁻¹ for IL-3 and 20 U ml⁻¹ for GM-CSF. An equivalent amount of non-immune normal rabbit serum (NRS) was used as a control. Each value represents arithmetic mean ± standard deviation for triplicate cultures.

further downstream we conclude that the *Csfm* gene of *op/op* mice cannot code for functional M-CSF protein. Although PCR amplification can generate random errors, it is unlikely to cause errors at the same position in four independent clones. The *Csfm* gene has recently been mapped and found to be close to the *op* locus on chromosome 3 (refs 5, 6), supporting our conclusion that the *op* locus is the *Csfm* gene itself.

In conclusion, the phenotypic abnormality in *op/op* mice is caused by the lack of a specific cytokine acting on haemopoietic cells. Such mutants, although severely defective in bone remodelling, are relatively healthy when maintained under pathogen-free conditions, suggesting that production of M-CSF is not necessary for the maintenance of various types of M-CSF responsive cells other than osteoclasts. The fact that *op/op* mice produced from matings of +/*op* heterozygotes appear normal at birth raises a question as to whether bone remodelling during early development is independent of M-CSF. This question is difficult to answer, because M-CSF is produced by maternal uterine epithelial cells directly above the decidua layer of placenta and is expected to be present in the tissues of *op/op* embryos^{17,18}. Thus, bone remodelling of the *op/op* fetus may proceed normally using maternal M-CSF, resulting in the normal development of the embryo. Further investigation of the pathophysiology of *op/op* mice, in the light of our present findings, will provide an insight into the role of M-CSF *in vivo* during postnatal life. □

Received 26 February; accepted 5 April 1990.

- Marks Jr. S. C. & Lane, P. W. *J. Heredity* **67**, 11-18 (1976).
- Wiktor-Jedrzejczak, W., Ahmed, A., Szczeplik, C. & Skelly, R. R. *J. exp. Med.* **156**, 1516-1527 (1982).
- Shultz, L. D. & Sidman, C. A. *Rev. Immun.* **5**, 367-403 (1987).
- Marks Jr. S. C., Seifert, M. F. & McGuire, J. L. *Metab. Bone Dis. & Rel. Res.* **5**, 183-186 (1984).
- Buchberg, A. M., Jenkins, A. A. & Copeland, N. G. *Genomics* **5**, 363-367 (1989).
- Gisselbrecht, S. et al. *Blood* **73**, 1742-1746 (1989).
- Sachs, L. *Science* **238**, 1374-1379 (1987).
- Metcalfe, D. *Nature* **339**, 27-30 (1989).
- Stanley, E. R. et al. *J. cell. Biochem.* **21**, 151-159 (1983).
- Saiki, R. K. et al. *Science* **239**, 487-491 (1988).
- Rajavashisth, T. B. et al. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1157-1161 (1987).
- Ladner, M. B. et al. *Proc. natn. Acad. Sci. U.S.A.* **85**, 6706-6710 (1988).
- Ladner, M. B. et al. *EMBO J.* **6**, 2693-2698 (1987).
- Takahashi, M. et al. *Biochem. biophys. Res. Commun.* **161**, 892-901 (1989).
- Ogawa, M. et al. *EMBO J.* **7**, 1337-1343 (1988).
- Nishikawa, S. I., Ogawa, M., Nishikawa, S., Kunisada, T. & Kodama, H. *Eur. J. Immun.* **18**, 1767-1771 (1988).
- Regenstreif, L. J. & Rossant, J. *dev Biol.* **133**, 284-294 (1989).

18. Pollard, J. W. *et al.* *Nature* **330**, 484–486 (1987).
 19. Kincade, P. W., Ralph, P. & Moore, M. A. S. *J. exp. Med.* **143**, 1265–1270 (1976).

ACKNOWLEDGEMENTS. This work is supported by the Ministry of Education, Science and Culture of Japan, the special aid programme of the Institute for Physics and Chemistry (RIKEN) and the NIH. We thank Chizu Furukawa for secretarial assistance.

Disruption of a C/EBP binding site in the factor IX promoter is associated with haemophilia B

Merlin Crossley & George G. Brownlee

Chemical Pathology Unit, Sir William Dunn School of Pathology, University of Oxford, Oxford OX1 3RE, UK

HAEMOPHILIA B (or Christmas disease) is an inherited, X-linked bleeding disorder caused by mutations in the gene for clotting factor IX. There is a rare class of patients, exemplified by haemophilia B Leyden¹, who suffer from haemophilia B as children but improve after puberty. In these patients, plasma factor IX concentrations are less than 10% of normal during childhood, but after puberty they gradually rise to between 40 and 80% of normal. Mutations clustered around the main transcription start point (defined as +1 (ref. 2)) have been reported in seven of these patients (at -20 (refs 1, 3, 4); -6 (refs 5, 6) and +13 (refs 7, 8)). To determine how these mutations interfere with factor IX expression, we have assayed for transcription factors binding to this area and have identified a nuclear factor-1 liver (NF1-L) binding site⁹ (-99 to -76) and a binding site for the CCAAT/enhancer binding protein (C/EBP)¹⁰ (+1 to +18). We show that the A → G mutation at +13 prevents the binding of C/EBP to this site. Furthermore, we show that C/EBP is capable of transactivating a cotransfected normal factor IX promoter but not the mutant promoter. This is the first natural mutation to be reported which disrupts a C/EBP binding site and is an illustration of the importance of this transcription factor in humans.

Factor IX, like the other vitamin K-dependent serine proteases involved in blood clotting, is expressed primarily in the liver¹¹. Therefore we used DNase I footprinting with rat liver nuclear extracts to analyse the region implicated by the point mutations. Two footprints were observed (Fig. 1). The factor binding to the distal element (-99 to -76) was heat-sensitive, the other (+1 to +18) was heat-stable (Fig. 1a). The distal footprint was observed using HeLa, human hepatoma cell lines (HepG2 and Hep3B (ref. 12)) and rat liver extracts, the proximal footprint only in rat liver extracts (Fig. 1 and our unpublished results). These observations together with a consideration of known consensus binding sites^{13–15}, led us to suspect that the distal footprint was caused by one of the nuclear factor-1 family of proteins and the proximal by C/EBP.

To test whether these proteins could indeed bind to these regions, we used an NF1-L fragment (containing N-terminal residues 4–240 and full DNA-binding activity⁹; gift from R. Cortese) and a C/EBP fragment (containing the 88 C-terminal residues and also capable of binding DNA¹⁶; gift from S. McKnight) in our binding assays. We show that these proteins do indeed bind to the expected sites and produce footprints essentially identical to those produced by nuclear extracts (Figs 1b and 2a).

As the A → G mutation at +13 (found in two unrelated individuals with Leyden-like haemophilia B^{7,8}) lies within the C/EBP binding site, we tested whether this mutation affected C/EBP binding. In our assays it abolished binding (Fig. 2b and c). This result was confirmed in competition experiments, where it was found that an oligonucleotide (residues -27 to +30) carrying the +13 mutation competed out the footprint roughly

100 times less efficiently than an equivalent normal oligonucleotide (data not shown).

The G → A mutation at -6 (also associated with the Leyden condition; ref. 5 and our unpublished results), is adjacent to the region protected by C/EBP (+1 to +18). Therefore it could affect C/EBP binding, either directly or indirectly, by generating a site for a second, unrelated DNA-binding protein which might itself prevent the binding of C/EBP. To investigate this possibility, a DNA fragment containing residues -189 to +21 and carrying the -6 mutation was used in footprinting experiments with rat liver extract. The -6 G → A mutation did not affect the proximal (C/EBP, +1 to +18) footprint or the distal (NF1-L, -99 to -76) footprint, nor was any new footprint seen (data not shown). This result indicates that the -6 mutation works by a different mechanism from that of the +13 mutation. Further evidence that this is the case is presented below.

Functional studies of the factor IX promoter are difficult, for several reasons. First, unlike most genes studied to date the

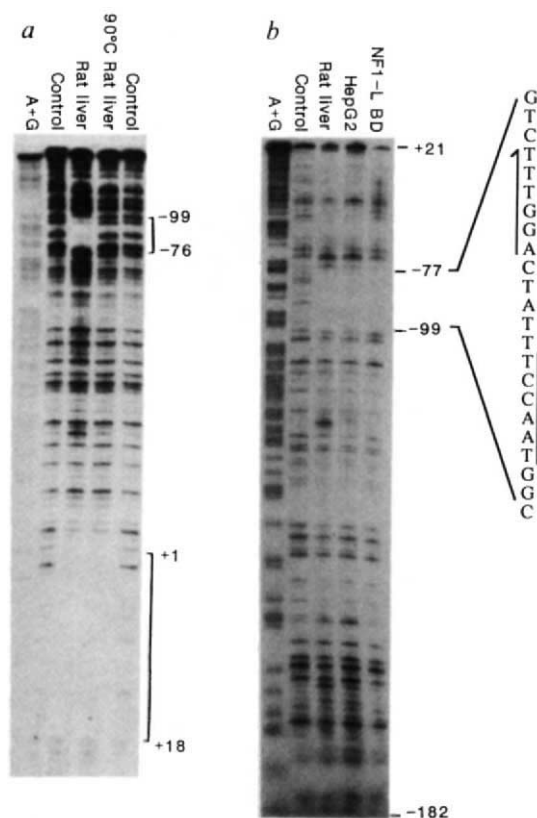
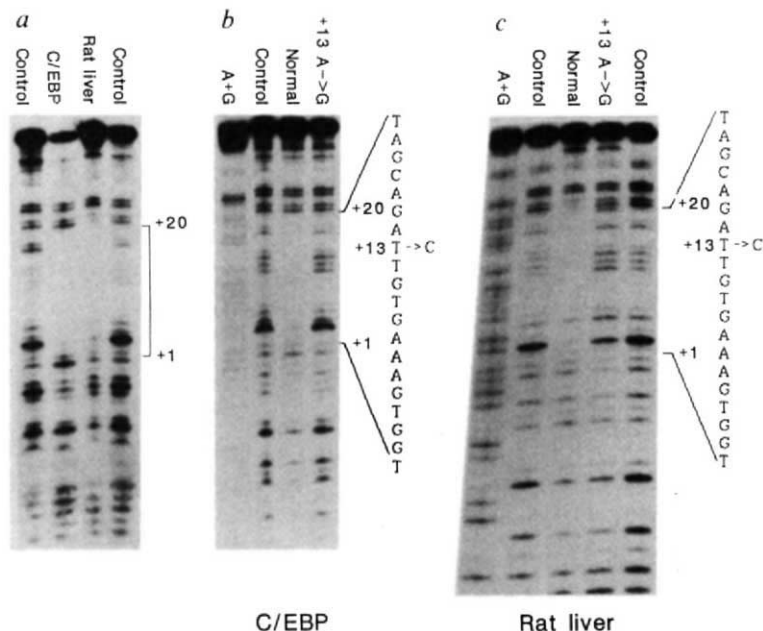


FIG. 1 DNase I footprinting of the factor IX promoter region. *a*, Restriction fragment (containing residues -189 to +21), ³²P end-labelled on the sense-strand and footprinted with either rat liver nuclear extract (20 µg), or a similar extract heated at 90 °C (10 min). Control lanes, throughout, contain no protein. A + G throughout refers to Maxam-Gilbert sequencing using formic acid. *b*, Same restriction fragment, end-labelled on the anti-sense strand and footprinted with rat liver nuclear extract (20 µg), HepG2 extract (20 µg) or purified NF1-L BD (NF1-L DNA-binding domain⁹ (60 ng)). The -189 to +21 fragment is derived from a *Hinf*I to *Nhe*I fragment, filled-in with Klenow and cloned into the *Sma*I site of pCAT00 (see Fig. 3 legend). To label the sense-strand, this fragment was cut with *Sau*3A, end-labelled using Klenow, excised by recutting with *Asp*718 and purified on a 6% non-denaturing polyacrylamide gel. The anti-sense strand was labelled similarly but was first cut with *Asp*718. Rat liver and HepG2 nuclear extracts were prepared by standard methods^{24–26}. End-labelled fragment (1–2 ng) was used in footprinting reactions (40 µl total volume) in 10 mM HEPES buffer, pH 7.8, 40 mM KCl, 3 mM MgCl₂, 4% Ficoll, 0.5 mM DTT and 2 µg poly(dI-dC). Proteins or extract were added first then, after 10 min at 20 °C, probe was added and after a further 10 min the mixture was digested (30–120 s) with 1 U RQ1 DNase (Promega). Samples were extracted with phenol-chloroform, precipitated with ethanol and analysed on 8% polyacrylamide-7M urea gels.

FIG. 2 DNase I footprinting of the proximal region, comparing normal and mutant probes. *a*, Footprinting of a restriction fragment (containing residues -27 to $+30$ and end-labelled on the anti-sense strand) with either purified C/EBP fragment (10 ng)¹⁶ or rat liver nuclear extract ($20\text{ }\mu\text{g}$). Controls contain no added protein. *b*, *c*, Footprinting of the same fragment and a similar fragment containing the $+13\text{ A}\rightarrow\text{G}$ mutation, with C/EBP fragment (10 ng) (*b*) or rat liver extract ($20\text{ }\mu\text{g}$) (*c*). The normal and mutant -27 to $+30$ fragments were obtained by cloning synthetic oligonucleotides into the *Sma*I site of pCATOO. End-labelling and footprinting reactions were as in Fig. 1 except that, where C/EBP was used, only 50 ng poly(dI-dC) was added to the footprinting mix. A + G, See legend to Fig. 1.



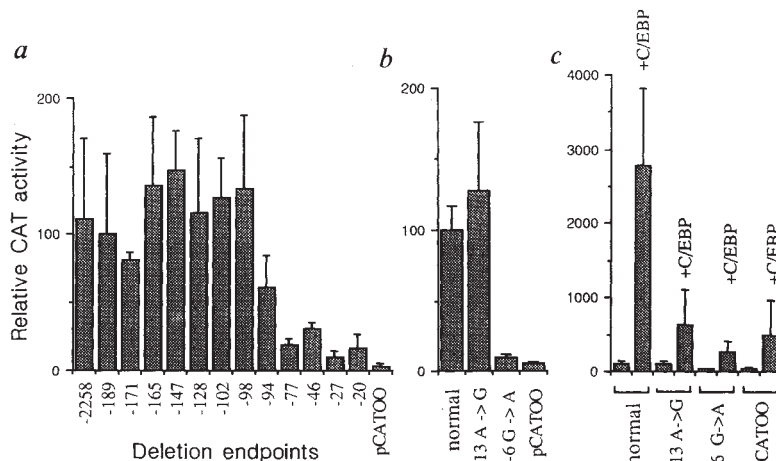
factor IX gene is expressed at low levels (the plasma concentration of factor IX is $5\text{ }\mu\text{g ml}^{-1}$ compared with that of albumin, $20\text{--}50\text{ mg ml}^{-1}$), consequently only very sensitive promoter assays are informative. Second, there is only one continuous cell line known to produce endogenous factor IX, namely HepG2. These cells do not produce detectable levels of secreted factor IX (ref. 12 and our unpublished results) but do produce small amounts of intracellular protein and messenger RNA¹⁷.

We linked a region presumed to contain the factor IX promoter ($-2,258$ to $+21$) to the reporter gene chloramphenicol acetyl transferase (CAT) and assayed for CAT activity in transiently-transfected HepG2, Hep3B and HeLa cells. We used a new, low-background vector (plasmid pCATOO), which contains two synthetic polyadenylation sites¹⁸ upstream of the CAT gene (Fig. 3). The region $-2,258$ to $+21$ is sufficient to produce low but detectable levels of CAT enzymatic activity in HepG2 cells. Only background levels of activity were observed in Hep3B and HeLa (unpublished results). A series of deletion

constructs were generated using *Bal*31 or convenient restriction sites and constructs were tested for activity in HepG2 cells (Fig. 3a). Deletions from $-2,258$ to -98 had no significant effect, but a further deletion to -77 caused a fivefold reduction in CAT activity. This result supports the hypothesis that the NF1-L binding site (-99 to -76) is important *in vivo*.

The $-6\text{ G}\rightarrow\text{A}$ and $+13\text{ A}\rightarrow\text{G}$ mutations were also tested in this functional assay. The -6 mutation reduced CAT activity to background levels but, surprisingly, the $+13$ mutation did not impair promoter activity (Fig. 3b). To test whether this was a consequence of the low concentrations of C/EBP in HepG2 cells¹⁹, we repeated the experiment in HepG2 cells simultaneously transfected with an expression plasmid for C/EBP (gift from S. McKnight¹⁹). C/EBP was capable of transactivating the normal promoter but it could not activate above background the promoter carrying the $+13\text{ A}\rightarrow\text{G}$ mutation (Fig. 3c). The background transactivation of pCATOO is most probably due to a cryptic C/EBP binding site(s) within the vector. Although

FIG. 3 Summary of the CAT assays in HepG2 cells. *a*, Relative activities of *Bal*31 deletion constructs. *b*, Comparison of normal promoter sequences (-189 to $+21$) with those carrying the $+13\text{ A}\rightarrow\text{G}$ and $-6\text{ G}\rightarrow\text{A}$ mutations. *c*, Comparison of the extents to which normal and mutant promoters are transactivated by C/EBP. The amount of CAT activity throughout is shown relative to plasmid pFIX-189 (see below), which is arbitrarily set at 100. The columns labelled 'normal' in (*b*) and (*c*) refer to this plasmid. All results are compiled from at least three independent experiments carried out with at least two independent preparations of plasmid. Standard errors around the mean are shown. The low-background CAT plasmid, pCATOO, was a modification of the previously described pUC119-based plasmid, pCATO (ref. 27; gift from P. Lamb). Plasmid pCATO was cut with *Sst*II, the site was blunt-ended and a *Bgl*II linker added. Two synthetic polyA sites¹⁸ (gifts from N. Proudfoot) were cloned into this site to form pCATOO. The promoter constructs were prepared as follows: pFIX-2,258 was made by cloning a blunt-ended *Nhe*I fragment ($-2,258$ to $+21$) into the *Sma*I site of pCATOO, which lies immediately upstream of the CAT gene; pFIX-189 was similarly constructed using a blunt-ended -189 to $+21$ *Hinf*I to *Nhe*I fragment. The remaining constructs were generated by *Bal*31-mediated deletion. HepG2 cells were grown on MEM medium supplemented with non-essential amino acids and 10% FCS. They were transfected²⁸ when 10–20% confluent. The β -galactosidase-encoding plasmids, pIRV²⁹ or pCH110³⁰ were used as



cotransfection controls. Extracts were prepared after 48 h and CAT and β -galactosidase activities measured³⁰. In *a* and *b*, CAT plasmid ($30\text{ }\mu\text{g}$) and pIRV ($2\text{ }\mu\text{g}$) were used. The transactivation experiment (*c*) was performed as described¹⁹. CAT plasmid ($5\text{ }\mu\text{g}$) and pCH110 ($2\text{ }\mu\text{g}$) were transfected together with either sonicated herring testes DNA ($15\text{ }\mu\text{g}$) or the C/EBP expression plasmid pMSV/C/EBP.

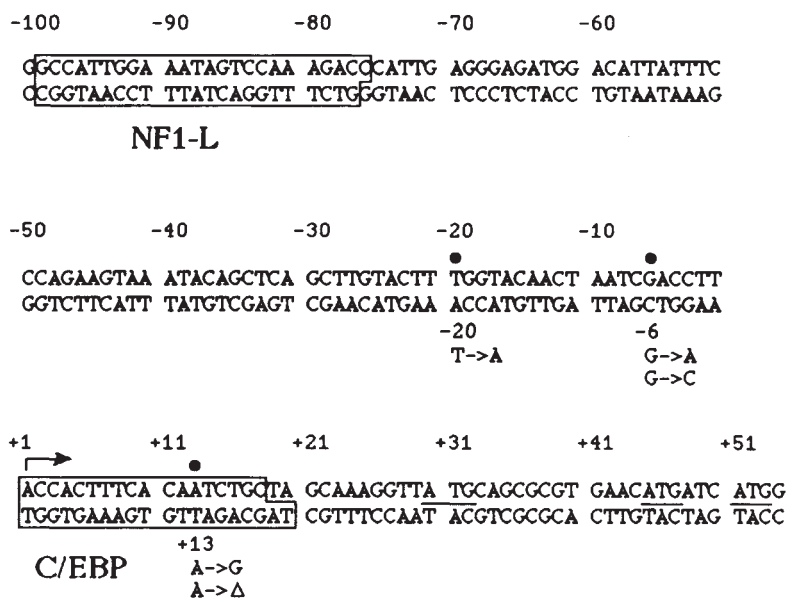


FIG. 4 Sequence of the 5' flanking region of the factor IX gene, containing the minimal promoter as defined by CAT assays in HepG2 cells. NF1-L and C/EBP binding sites are boxed. The primary start point of transcription² is marked with an arrow and potential translation-initiation codons are underlined. Previously described mutations known to be associated with the Leyden form of haemophilia B are indicated by dots (-20 T→A (refs 1, 3, 4); -6 G→A (ref. 5) and our unpublished results; -6 G→C (ref. 6); +13 A→G (refs 7, 8); and the single base deletion (Δ) at +13 (ref. 7).

such transactivation experiments may not reflect the *in vivo* situation in quantitative terms, these results clearly demonstrate that C/EBP can activate the factor IX promoter and that the +13 mutation interferes with this activation. Furthermore, the fact that the +13 mutation impairs promoter function only when tested in HepG2 cells transfected with a C/EBP encoding plasmid, whereas the -6 mutation reduces promoter activity in both systems (Fig. 3b and c), supports our assertion that the -6 and +13 mutations exert their effects through different mechanisms.

In conclusion, we have identified an NF1-L (-99 to -76) and a C/EBP binding site (+1 to +18) in the 5' flanking region of the human factor IX gene (Fig. 4). The results obtained from *in vivo* assays of promoter activity, carried out in HepG2 cells, are consistent with the *in vitro* footprinting results. Deletion of the NF1-L binding site caused a fivefold reduction in CAT activity. The +13 A→G mutation, which is associated with haemophilia B^{7,8} and disrupts a C/EBP binding site *in vitro*, did not impair promoter activity in ordinary HepG2 cells but did in HepG2 cells cotransfected with a C/EBP expression plasmid. The -6 G→A mutation associated with haemophilia B caused a significant reduction in CAT activity in our assays. The mutation did not influence the binding of NF1-L or C/EBP in footprinting experiments. It is probable that as-yet-uncharacterized protein(s) bind to the -6/-20 region(s).

Although we have shown that there are NF1-L and C/EBP binding sites within the minimal promoter, as defined in HepG2 cells, and that C/EBP can transactivate the promoter, it is possible that other proteins may also bind to these sites and be important for transcription *in vivo*. Other members of the NF1 family²⁰ may bind. Moreover it has recently been shown that C/EBP is one of a family of proteins, encoded by separate genes (Z. Cao and S. McKnight; P. Johnson, personal communication). These proteins possess closely related DNA-binding domains and may also be capable of interacting with the factor IX promoter.

Although C/EBP binding sites have been found both upstream and downstream of many transcription start sites¹⁵, the finding of a binding site so close to the start site is unusual. The significance of this finding, in relation to the role of C/EBP during the initiation and process of transcription remains to be defined.

The significant increase in factor IX expression in haemophilia B Leyden patients after puberty is not understood, although it is directly or indirectly due to increasing levels of

testosterone^{4,21}. We feel it is best explained by postulating the existence of a nearby androgen-responsive element, which contributes to promoter strength in these patients after puberty⁷. We are aware of no evidence that the promoter is responsive to androgen in the normal population.

Naturally occurring promoter mutations are a rare and promising source of information on mammalian gene expression and transcription factors²². C/EBP in particular may be important in fully differentiated, non-proliferating cells^{19,23}. The finding that the disruption of a C/EBP binding site is associated with haemophilia B in humans supports the view that C/EBP (or, more correctly, the human analogue of the rat protein C/EBP) plays an important part in gene expression in the liver *in vivo*. □

Received 7 February; accepted 6 April 1990.

- Veltkamp, J. J., Meilof, J., Remmelts, H. G., Van der Vliet, D. & Loeliger, E. A. *Scand. J. Haematol.* **7**, 82-90 (1970).
- Anson, D. S. *et al.* *EMBO J.* **3**, 1053-1060 (1984).
- Briët, E., Bertina, R. M., Van Tilburg, N. H. & Veltkamp, J. J. *New Engl. J. Med.* **306**, 788-790 (1982).
- Reitsma, P. H., Bertina, R. M., Ploos van Amstel, J. K., Riemens, A. & Briët, E. *Blood* **72**, 1074-1076 (1988).
- Kurachi, K., Hirose, S., Fahner, J. B., Wu, C. T. & Salier, J. P. *Thromb. Haemostasis* **62**, 154 (1989).
- Attree, O., Vidau, M., Vidau, D., Laverne, J. M. & Goossens, M. *Thromb. Haemostasis* **62**, 8 (1989).
- Reitsma, P. H., Mandala, T., Kasper, C. K., Bertina, R. M. & Briët, E. *Blood* **73**, 743-746 (1989).
- Crossley, P. M., Winship, P. R., Black, A., Rizza, C. R. & Brownlee, G. G. *Lancet* **i**, 960 (1989).
- Paonessa, G., Gounari, F., Frank, R. & Cortese, R. *EMBO J.* **7**, 3115-3123 (1988).
- Landschulz, W. H., Johnson, P. F., Adashi, E. Y., Graves, B. J. & McKnight, S. L. *Genes Dev.* **2**, 786-800 (1988).
- Fair, D. S. & Bahnak, B. R. *Blood* **64**, 194-204 (1984).
- Aden, D. P., Fogel, A., Plotkin, S., Damjanov, I. & Knowles, B. B. *Nature* **282**, 615-616 (1979).
- Nowack, J., Borgmeyer, U., Püschel, A. W., Rupp, R. A. W. & Sippel, A. E. *Nucleic Acids Res.* **13**, 2045-2061 (1985).
- Gronostajski, R. M. *Nucleic Acids Res.* **15**, 5545-5559 (1987).
- Ryden, T. A. & Beemon, K. *Molec. cell. Biol.* **9**, 1155-1164 (1989).
- Vinson, C. R., Sigler, P. B. & McKnight, S. L. *Science* **246**, 911-916 (1989).
- Huang, M.-N., Smith, K., Evans, J. P., Wolde, S. T. & High, K. A. *Thromb. Haemostasis* **62**, 81 (1989).
- Levitt, N., Briggs, D., Gil, A. & Proudfoot, N. J. *Genes Dev.* **3**, 1019-1025 (1989).
- Friedman, A. D., Landschulz, W. H. & McKnight, S. L. *Genes Dev.* **3**, 1314-1322 (1989).
- Santoro, C., Mermod, N., Andrews, P. C. & Tjian, R. *Nature* **324**, 218-224 (1988).
- Briët, E., Wijnands, M. C. & Veltkamp, J. J. *Ann. Intern. Med.* **103**, 225 (1985).
- Martin, D. I. K., Tsai, S.-F. & Orkin, S. H. *Nature* **338**, 435-438 (1989).
- Birkenmeier, E. H. *et al.* *Genes Dev.* **3**, 1146-1156 (1989).
- Graves, B. J., Johnson, P. F. & McKnight, S. L. *Cell* **44**, 565-576 (1986).
- Gorski, K., Carneiro, M. & Schibler, U. *Cell* **47**, 767-776 (1986).
- Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. *Nucleic Acids Res.* **11**, 1475-1489 (1983).
- Lamb, P., Watt, P. & Proudfoot, N. J. *Progress in Clinical and Biological Research, Haemoglobin Switching* Vol. **316**, 269-278 (Liss, New York, 1989).
- Graham, F. L. & van der Eb, A. J. *Virology* **52**, 456-467 (1973).
- Bedington, R. S. P., Morgernstern, J., Land, H. & Hogan, A. *Development* **106**, 37-46 (1989).
- Herbomel, P., Bourachot, B. & Yaniv, M. *Cell* **39**, 653-662 (1984).

ACKNOWLEDGEMENTS. We thank Peter Lamb, Emma Whitelaw and Peter Winship for helpful discussions and Dr Charles Rizza for clinical information. M.C. is supported by the Rhodes Trust.

Morphological transformation of human keratinocytes expressing the LMP gene of Epstein-Barr virus

Robin Fåhræus*, Lars Rymo†, John S. Rhim‡ & George Klein*

* Department of Tumor Biology, Karolinska Institutet, Box 60400, S-104 01 Stockholm, Sweden

† Department of Medical Chemistry, Gothenburg University, S-400 33 Gothenburg, Sweden

‡ Laboratory of Cellular and Molecular Biology, National Cancer Institute, Bethesda, Maryland 20892, USA

THE association of Epstein-Barr virus (EBV) with nasopharyngeal carcinoma (NPC) has been known for some time^{1,2}, but the precise role of EBV in this cancer is poorly understood, due partly to the lack of an *in vitro* system for studying NPC cells and the effect of EBV on epithelial cells. Biopsies of NPC tumours have revealed expression of the EBV latent membrane protein (LMP) in 65% of cases³, suggesting that in at least some NPC tumours LMP may contribute to cell transformation. Here we address the question of the effect of LMP expression on epithelial cells. Transfection of an immortalized, non-tumorigenic keratinocyte cell line (RHEK-1) with the LMP gene causes a striking morphological transformation: the originally flat, polygonal colonies change to bundles of spindle-shaped cells that form multilayer foci, and

cytokeratin expression is down-regulated. Our results suggest that LMP expression may be an important causal factor in the development of NPC.

RHEK-1 cells were transfected with either plasmid pJ132 or pJ144, which carry the genes for LMP and EBNA-2 (EBV nuclear antigen) respectively and the guanine phosphoribosyl transferase selectable marker. After selection on medium containing mycophenolic acid and xanthine, cells were analysed for the expression of EBNA-2 and LMP respectively (Fig. 1). A band of relative molecular mass (M_r) 90,000 (90K), corresponding to the expected EBNA-2 size, was detected in the 100% positive EBNA-2-transfected clones, but not in the LMP-transfected or the control cells. Six different LMP-transfected clones all contained a new 63K band, detected by the anti-LMP monoclonal antibody, S-12. This band corresponded in size, to the known LMP band of EBV-transformed lymphoblastoid cell line, IB-4, and was not detected in the EBNA-2 transfectant or in the EBV-negative control line.

PKK-1 and PKK-2 are monoclonal antibodies against cytokeratins. PKK-1 reacts against cytokeratins 19, 18, 8 and PKK-2 against 19, 17, 16 and 7 (ref. 4). AE1/AE3 is a pooled mixture of antibodies that reacts against cytokeratin antigens of M_r values between 40 and 67K (ref. 5). PKK-1 and PKK-2 reacted against a band of approximately 40K in the control and in EBNA-2-transfected RHEK-1 cells. In addition two weak bands at approximately 90–95K not corresponding to any known cytokeratins were detected with the PKK-2 antibody in RHEK-1 and in cells transfected with EBNA-2. Neither of these bands appeared in the LMP-transfected cells (Fig. 2a, b). The AE1/AE3 mixture reacted against a variety of bands between approximately 40 and 66K in the control cells. A similar but

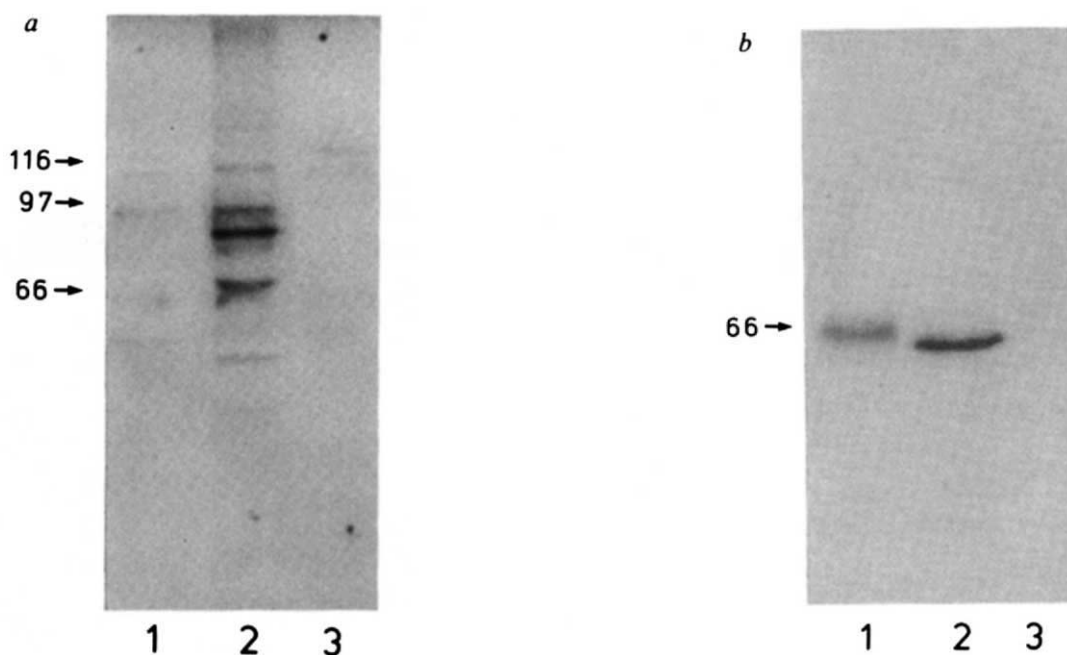


FIG. 1 *a*, RHEK-1 cells transfected with cloned EBNA-2, 100% positive by immunofluorescence staining, were incubated on immunoblots with the human polyvalent EBV-positive serum PG. A new band at 90K, corresponding to EBNA-2 in the B95-8 cells (lane 2), was detected (lane 1), but not in the untransfected RHEK-1 cells (lane 3). *b*, LMP-specific monoclonal antibody, S-12, detected a novel 63K band in the LMP-transfected RHEK-1 cells (lane 1) but not in RHEK-1 cells (lane 3). This band corresponded to the LMP of the EBV-transformed lymphoblastoid cell line, IB4, which was running slightly faster (lane 2).

METHODS. Cells were trypsinized and washed once in PBS before being sonicated for 3×10 s in a 50 mM sodium phosphate buffer pH 6.8 containing 2% SDS, 2 mM DTT, 5% glycerol and a pinch of bromophenol blue. The lysate was boiled for 5 min and applied to a 7.5% SDS-PAGE gel essentially as

described by Laemmli¹⁸. After separation the proteins were routinely transferred to a nitrocellulose filter in a Bio-Rad mini-blot apparatus. The filters were stained with Ponceau S (Sigma) and the position of the molecular weight markers (Bio-Rad) was noted before they were destained and pre-absorbed in PBS containing 5% non-fat dry milk. They were then incubated with serum or monoclonal antibodies overnight at 4 °C and exposed to antibodies conjugated to anti-IgG alkaline phosphatase antibodies conjugated to (Bio-Rad) for 2 h. After six washes in 0.5% Tween 20, the phosphatase substrate was added. IB4 and the EBV-negative B-cell lymphoma line BJAB were used as positive and negative controls respectively for the detection of EBNA-2 and LMP by immunoblotting (not shown). The S-12 monoclonal antibody directed against the C-terminal of LMP and a polyclonal EBV-positive human serum, PG were used for identifying LMP and EBNA-2.

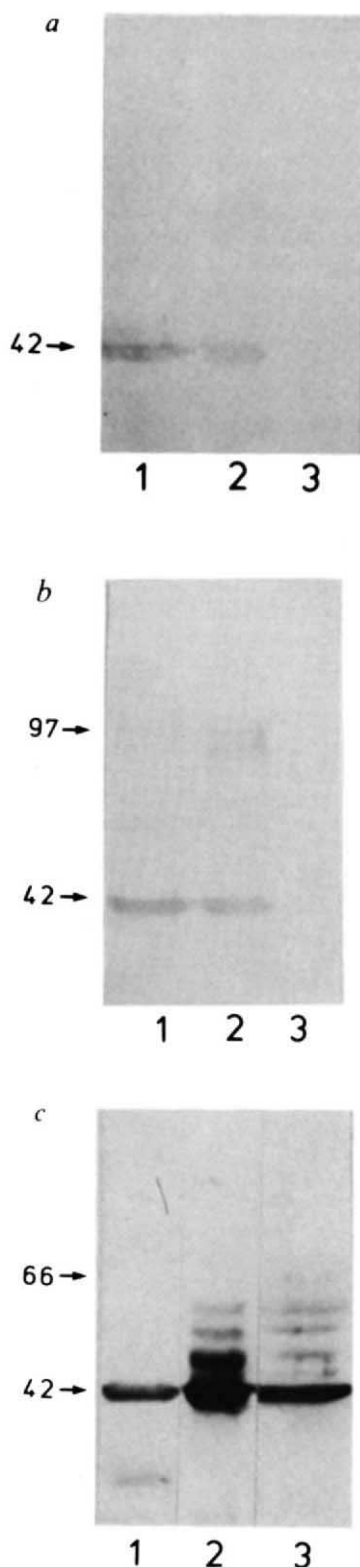


FIG. 2 Immunoblotting with the monoclonal antibodies PKK-1 (a) and PKK-2 (b) (Labsystem) against cytokeratins 19, 18 and 8 and against 19, 17, 16 and 7 respectively. Lane 3 shows the LMP-transfected RHEK-1 cells compared with RHEK-1 in lane 1 and the RHEK-1 cells transfected with EBNA-2 in lane 2. c, Pattern of cytokeratins between 40 and 66K detected in the RHEK-1 (lane 2) and the EBNA-2-transfected RHEK-1 cells (lane 3) with the AE1/AE3 (Boehringer & Mann) pool of mixed monoclonal antibodies. The LMP-transfected RHEK-1 cells (lane 1) shows a down-regulation of cytokeratin expression.

METHODS. The immunoblotting is describing under Fig. 1. All antibodies were used in a dilution 1:500.

weaker pattern of bands also appeared in the cells transfected with EBNA-2. The same antibody mixture reacted against only one band of about 42K in LMP-transfected cells (Fig. 2c). Anticomplementary immunofluorescence (ACIF) with EBNA-2-specific antibodies gave a brilliant, finely dispersed nuclear fluorescence with EBNA-2-transfected, but not LMP-transfected or control cells (Fig. 3).

The LMP-transfected cells changed to a more fibroblastic morphology with altered architecture and growth pattern. The flat and even cobblestone-like colonies that characterized keratinocytes cultured *in vitro* could not be recognized. Instead, the cells grew in tight, disorganized bundles that became partly multilayered and formed foci long before they reached confluence. The EBNA-2-transfected cells showed no detectable change compared to the untransfected cells. A similar change to fibroblastic morphology and piling-up in foci was previously observed in the same line after transfection with *v-ras* (ref. 6).

The EBV genome has been detected in the oropharynx and salivary glands⁷ and virus-producing foci have been identified in the oral hairy leukoplakia of AIDS patients⁸, but no epithelial tumors other than NPC are associated with EBV⁹. Six nuclear antigens (EBNA-1-6) and LMP are regularly expressed in virally transformed immortalized B-cell lines (LCL)¹⁰. Their function is not, or only incompletely, known. EBNA-1 is expressed in all EBV-infected cells. It binds to the origin of replication (*oriP*) and is required for maintenance of multiple viral genomes in an episomal form¹¹. EBNA-2 is essential for B-cell activation, a prerequisite for B-cell immortalization¹². The receptor-like structure of LMP has promoted speculation on its possible involvement in an autocrine circuit¹³ and can transform established rodent fibroblast lines *in vitro*^{14,15}. The transforming effect and the down-regulation of keratin expression in the LMP-transfected keratinocytes, together with the fact that LMP is only expressed in Burkitt's lymphoma after long periods of propagation *in vitro* and in parallel with a drift toward a more

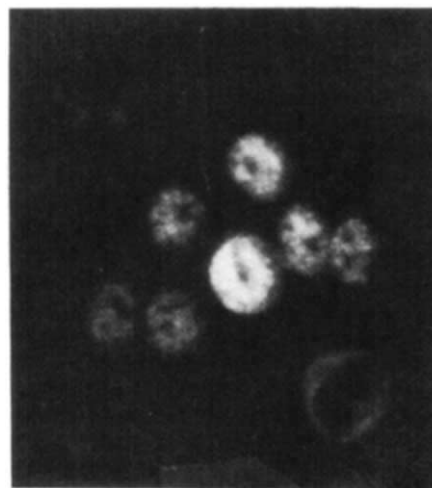


FIG. 3 Nuclear fluorescence of EBNA-2-transfected, uncloned RHEK-1 cells stained against an EBV-positive serum, pre-absorbed with EBNA-2-negative P3HR-1 cells.

METHODS. The RHEK-1/EBNA-2 cells were cultured in a tissue-chamber slide (LabTech) and fixed in -20°C acetone/methanol (2/1) for 15 min. Three EBNA-2-antibody-negative and three EBNA-2-positive polyclonal sera were used. One of the latter (KF) was absorbed with P3HR-1 cells positive for EBNA-1, EBNA-3 and LMP and negative for EBNA-2. After incubation for 30 min at 37°C in a wet chamber EBV-negative human complement was added and incubated at 37°C for 20 min. Anti-B1C-FITC (DAKO) was added for 30 min at room temperature. The washed cells were counterstained for 5 min in Evans Blue.

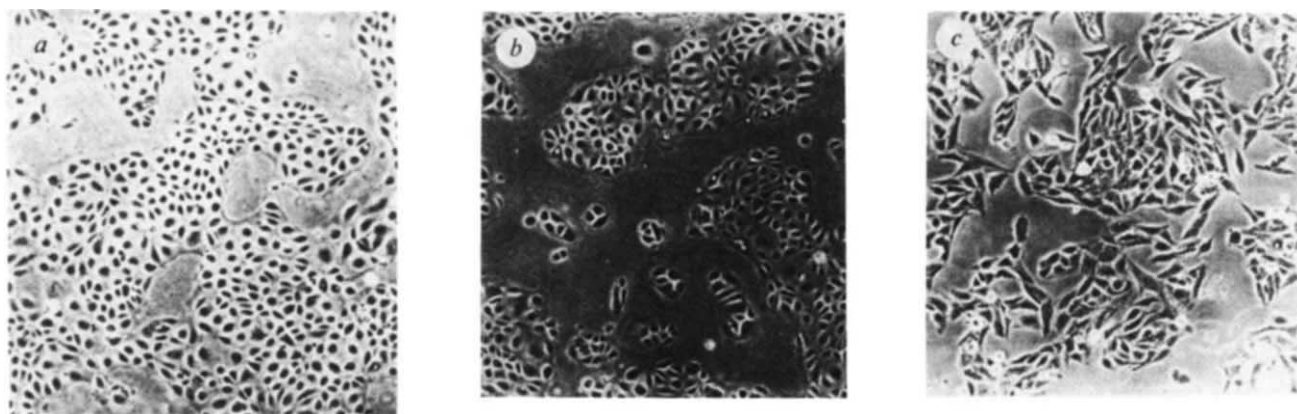


FIG. 4 *a*, RHEK-1 cells: primary foreskin keratinocytes immortalized with Ad12-SV40 hybrid virus. *b*, RHEK-1 cells transfected with the EBNA-2 gene, and showing no change in morphology. *c*, RHEK-1 cells transfected with the LMP gene. Note the change in morphology and growth in spindle-shaped bundles.

METHODS. Growth and maintenance medium consisted of RPMI with 10% fetal calf serum, hydrocortisone ($2 \mu\text{g ml}^{-1}$), penicillin and streptomycin. Cells transfected with EBNA-2 and LMP were kept on a selective medium containing mycophenolic acid (MPA) ($10 \mu\text{g ml}^{-1}$) and xanthine ($100 \mu\text{g ml}^{-1}$). Plasmids pJ132 and pJ144 carry the LMP and EBNA-2 genes, respectively, under the control of the SV40 early promoter. They were made by cloning a *Bgl*II-*Bam*HI fragment containing the SV40 enhancer and promoter sequence into the unique *Bam*HI site of plasmid pSV2gpt/*Bgl*II (ref. 19) resulting in plasmid pJ124 which allows transcription from both SV40

promoters in the same direction. pJ132 was made by cloning the LMP-encoding, 3,535-bp *Bam*HI-*Sac*II fragment of B95-8 EBV DNA²⁰ into the *Bam*HI site of pJ124 using *Bam*HI linkers. pJ144 was constructed by cloning the *Bgl*II fragment of pE-A6, encoding EBNA-2²¹ into the *Bam*HI site of pJ124. The DNA was transfected by electroporation²². Cells were trypsinized and washed with PBS. DNA was added ($10 \mu\text{g}$ per 5×10^6 cells) directly to the resuspended cell pellet together with $100 \mu\text{l}$ of HeBS buffer. The mixture was transferred into cuvettes, electroporated with 170 V at $960 \mu\text{F}$ and put on ice for 10 min. The cells were cultured in regular medium and selection started after 24 h by adding $5 \mu\text{g ml}^{-1}$ MPA and $50 \mu\text{g ml}^{-1}$ xanthine. The concentration of MPA and xanthine was doubled after a week and maintained subsequently at the same level. The cells were then cloned in a multichamber dish and tested for either LMP or EBNA-2 expressions.

LCL-like phenotype¹⁶, raises the question of a role for LMP in the aetiology of NPC, a notoriously little differentiated or anaplastic tumour. This attractive hypothesis must be reconciled with the fact, however, that LMP is not detected in approximately 35% of NPC tumour biopsies³. The possibility that this could be due to technical artefacts involved in detecting LMP in tumour tissue must first be excluded. If the finding is real it may relate to the documented fact that LMP is highly immunogenic for T cells¹⁷. It is conceivable that LMP-negative tumours may represent a secondary immunoselected variant. Alternatively, LMP might not be essential for the genesis of the tumour but could provide an advantage for its progressions. □

Presentation of viral antigen controlled by a gene in the major histocompatibility complex

V. Cerundolo*, J. Alexander†, K. Anderson†, C. Lamb†, P. Cresswell†, A. McMichael*, F. Gotch* & A. Townsend*

* Institute of Molecular Medicine, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK

† Department of Microbiology and Immunology, Duke University Medical Centre, Durham, North Carolina 27710, USA

WE describe a mutant human cell line (LBL 721.174)^{1,2} that has lost a function required for presentation of intracellular viral antigens with class I molecules of the major histocompatibility complex (MHC), but retains the capacity to present defined epitopes as extracellular peptides. The cell also has a defect in the assembly and expression of class I MHC molecules^{2,3,4}, which we show can be restored by exposure of the cells to a peptide epitope. This phenotype suggests a defect in the association of intracellular antigen with class I molecules similar to that described for the murine mutant RMA-S (ref. 5), but in the present case the genetic defect can be mapped within the MHC locus on human chromosome 6.

The human mutant cell line LBL 721.174 (0.174) was selected from the B-cell line LBL 721 by exposure to gamma rays and sequential treatments with monoclonal antibodies against class I and class II molecules and complement¹. It has a defect in assembly and transport of class I molecules that has resulted in complete loss of expression of B5, and 70–80% reduction in A2, at the cell surface^{2,4}. These features resemble the phenotype of the murine mutant RMA-S⁵, which has a selective defect in presentation of intracellular viral antigens to class I-restricted

Received 25 January; accepted 27 March 1990.

- Klein, G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **71**, 4737–4741 (1974).
- Klein, G. in *The Epstein-Barr Virus* (eds Epstein, M. A. & Achong, B. G.) 339–350 (Springer, Berlin, 1979).
- Fähraeus, R. *et al.* *Int. J. Cancer* **42**, 329–338 (1988).
- Moll, R. *et al.* *Cell* **31**, 11 (1982).
- Sun, T.-T. *et al.* in *The Cancer Cells vol 1* (eds Levin, A., Vaude Woude, G. F., Topp, W. & Watson, J.) 169–176 (Cold Spring Harbor, New York, 1984).
- Rhim, J. S., Gilbert, J., Arnstein, P., Price, F., Sanford, K. & Aaronson, S. *Science* **227**, 1250–1252 (1985).
- Sixbey, J. W., Nedrud, J. G., Raub-Traub, N., Hanes, R. A. & Pagano, J. S. *N. Engl. J. Med.* **310**, 1225–1230 (1984).
- Greenspan, J. S. *et al.* *N. Engl. J. Med.* **313**, 1654–1671 (1985).
- Andersson-Anvret, M., Forsby, N., Klein, G. & Henle, W. *Int. J. Cancer* **20**, 486–494 (1977).
- Dillner, J. & Kallin, B. *Adv. Cancer Res.* **50**, 95–158 (1988).
- Yates, J., Warren, N., Reisman, D. & Sugden, B. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3806–3810 (1981).
- Skare, J. *et al.* *J. Virol.* **55**, 268–297 (1985).
- Mann, K. P. & Thorley-Lawson, D. *J. Virol.* **61**, 2100–2108 (1987).
- Wang, D., Liebowitz, D. & Kieff, E. *J. Virol.* **62**, 2337–2346 (1988).
- Baichwald, V. R. & Sugden, B. *Oncogene* **2**, 461–467 (1988).
- Rowe, M. *et al.* *EMBO J.* **6**, 2743–2751 (1987).
- Murray, R. *et al.* *J. Virol.* **62**, 3747–3755 (1988).
- Laemmli, U. K. *Nature* **227**, 680–685 (1970).
- Rymo, L., Klein, G. & Ricksten, A. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3435–3439 (1985).
- Arrand, J. R. *et al.* *Nucleic Acids Res.* **9**, 2999–3014 (1981).
- Ricksten, A., Svensson, C., Weinder, C. & Rymo, L. *J. gen. Virol.* **68**, 2407–2418 (1987).
- Potter, H., Weir, L. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7161–7165 (1984).

ACKNOWLEDGEMENTS. We thank Ann Karlsson for technical assistance and C.-J. Dalsgaard's group (Department of Anatomy, Karolinska Institutet) for help and advice. This investigation was supported by Grants from the NCI, and from the Swedish Medical Research Council and the Swedish Cancer Society.

T cells. We therefore looked for defects in antigen presentation in 0.174 cells.

We studied the ability of 0.174 cells to present an epitope (amino acids 58–68) of influenza matrix protein to a specific A2-restricted T-cell line (JM22) (ref. 6). The parental cells were lysed efficiently after infection with either a recombinant vaccinia virus that makes matrix protein (Fig. 1b), or influenza virus (data not shown). In contrast, the mutant 0.174 cells were not recognized in identical conditions (Fig. 1a). The lack of recognition of the infected 0.174 cells by JM22 T cells was not due to inefficient infection, as 0.174 and LBL 721 synthesized similar amounts of viral proteins after infection (data not shown).

Mutant cells, despite their reduced expression of A2 at the cell surface, required roughly 100-fold less peptide epitope than the wild-type cell to sensitize them for lysis by JM22 cytotoxic T lymphocytes (CTL) (Fig. 1a–c). These results showed that the defect in presentation in 0.174 was selective for intracellular antigen, and revealed an increased capacity of the mutant cell to present extracellular peptide.

To confirm that the A2 molecules of 0.174 were functional, and to establish that the mutation responsible for the loss of antigen presentation is on chromosome 6, we compared two derivatives of 0.174 obtained by fusion to the T lymphoma CEM⁷. The 0.174×CEM hybrid (T1) expressed high levels of A2 and B5 derived from 0.174⁷. Its derivative (T2), selected for loss of both copies of human chromosome 6 provided by CEM, has the low class I expression phenotype of 0.174. This indicated that a *trans*-acting element encoded on CEM chromosome 6 enhanced class I expression in T1⁷.

We found that T1 presented intracellular antigen, but T2 did not (Fig. 1d, e). These results placed the mutation responsible for the defect in antigen presentation in 0.174 cells on human chromosome 6. In addition they confirmed that this defect is

not due to a mutation in A2 molecules. The mutant 0.174 bears a large homozygous deletion on the short arm of chromosome 6 between the DP alpha 2 locus and the complement gene cluster⁸. It is therefore likely that the gene(s) responsible for the defect in antigen presentation exist within this deleted region of the MHC.

To ensure that the defect in intracellular antigen presentation in the mutant was not limited to A2 and influenza matrix protein, T2 cells were transfected with several other class I genes and tested for presentation of epitopes of influenza virus that we have defined with synthetic peptides. In each case presentation of intracellular antigen was lost, whereas extracellular peptides were presented efficiently (table 1).

These findings show that the genetic defect in mutant 0.174 cells, like that in the murine mutant RMA-S⁵, results in both loss of presentation of intracellular antigens and defective expression of class I molecules at the cell surface. These two traits could be consistent with a single gene defect by postulating that class I assembly and transport is facilitated by association with peptides derived from intracellular antigens⁵.

The residual expression of A2 on 0.174, and the enhanced ability of the cells to present extracellular peptides, could be accounted for by assuming that a low level of assembly in the endoplasmic reticulum (ER), and expression of 'empty' A2 molecules at the cell surface, can occur in the absence of endogenous peptides. Murine class I glycoproteins encoded by transfected genes are similarly expressed at the cell surface in T2^{9,10}, but also do not present endogenous antigens to CTL (for example K^b in Table I).

By analogy with RMA-S⁵, exposure of 0.174 or T2 cells to a sufficiently high concentration of extracellular peptides should enhance surface expression of class I molecules. Figure 2 shows that A2 expression at the 0.174 cell surface increased up to threefold after exposure to the peptides M58–68 or K62 (a more

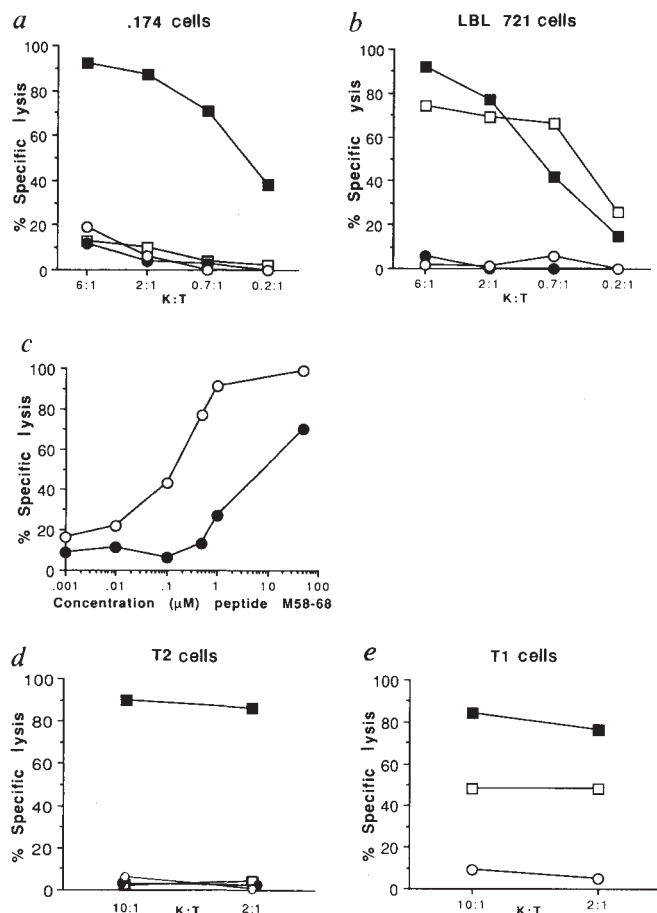


FIG. 1 Antigen presentation by 174 cells. *a, b*, JM22 T-cell line specific for the influenza matrix protein epitope 58–68 recognizes 0.174 cells treated with peptide M58–68, but not infected with a recombinant vaccinia which makes matrix protein (M1-VAC). Target cells were: *a*, mutant 0.174, *b*, the wild-type LBL 721. Target cells were: □, infected with M1-vac; ■, treated with peptide M58–68; ●, infected with wild-type vaccinia virus; ○, untreated. *c*, The mutant 0.174 cells present peptide M58–68 more efficiently than wild-type LBL 721 cells. Target cells were: ○, 0.174 and ●, LBL 721 cells exposed to the peptide M58–68 at the concentrations shown. *d, e*, Fusion of 0.174 with CEM cells restores presentation of intracellular antigen. In *d*, target cells were T2 cells (0.174×CEM selected for loss of CEM chromosomes 6); and in *e*, T1 (0.174×CEM). Target cells were: □, T2 cells infected with M1-vac; ■, T2 cells treated with peptide M58–68; ●, T2 cells infected with wild-type vaccinia virus; ○, T2 cells untreated. In *(e)*, target cells were: □, T1 cells infected with influenza virus X31; ■, T1 cells treated with peptide M58–68; ○, T1 cells untreated.

METHODS. JM22 is an influenza A virus-specific A2-restricted T-cell line which recognizes specifically the matrix 1 epitope 58–68 isolated as described⁶. Its activity was tested by Cr⁵¹ release assay on peptide treated (5×10^{-5} M for 1 h) or M1-vac infected (5 plaque-forming units per cell) or influenza X31 virus infected target cells⁶. In *c*, target cells were exposed to various concentrations of peptide M58–68 for 1 h, then washed twice and exposed to JM22 plated at the killer to target ratio (K:T) of 10:1. Cr⁵¹-labelled cells were plated at 5×10^3 cells per well at the K:T ratios shown. Supernatants were collected after 4 h and specific lysis calculated as described¹⁶.

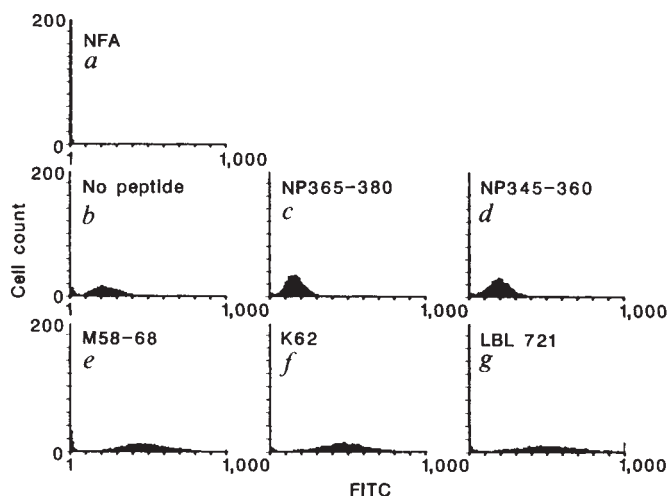


FIG. 2 Induction of A2 surface expression on 0.174 cells by peptide M58-68. *a*, Background staining 0.174 cells with no first antibody (NFA); *b*, basal level MA2.1-specific staining detected on untreated 0.174 cells; *c-f*, 0.174 cells stained with MA2.1 after treatment with the peptides shown; *g*, MA2.1-specific staining detected on wild type LBL 721 cells.

METHODS. Cells (5×10^5) were exposed to medium alone (RPMI/10% FCS) or medium containing peptides at the concentration of 10^{-4} M in a total volume of 1 ml for 12 h at 37 °C, then collected and stained by indirect immunofluorescence⁵. The first layer was MA2.1 (protein A-Sepharose purified and used at $5 \mu\text{g ml}^{-1}$), and the second layer was FITC-labelled affinity-purified goat anti-mouse antibody (Sigma) at 1:60 dilution. Samples were analysed on an Ortho Cytofluorograf.

FIG. 3 Effect of peptide on conformation of A2 molecules detected by the MA2.1 antibody. *a*, Peptide induces a conformational change in the alpha-1 domain of the A2 heavy chain in 0.174 and T2 cells. Lane A, untreated 0.174 cells; B, 0.174 cells + NP345-360 (negative control); C, 0.174 cells + M58-68; D, 0.174 cells + K62; E, untreated LBL721 cells; F, untreated T2 cells; G, T2 cells + NP 345-360; H, T2 cells + M58-68; I, T2 cells + K62; J, untreated T1 cells. *b*, Peptide induces the conformational change in endo H-sensitive heavy chains. Samples were immunoprecipitated with MA2.1. Lanes A-D, untreated T2 cells; lanes E-H T2 cells treated with peptide NP 345-360; lanes I-L, T2 cells treated with peptide M58-68. Immunoprecipitated A2 molecules in lanes B, D, F, H, J, L were digested with endo H, and in lanes A, C, E, G, I, K, they were mock-digested. The position after digestion of endo H-resistant (R) and -sensitive (S) A2 molecules are marked on the right. Samples in lanes A, B, E, F, I, J were lysed after 15-min labelling, samples in lanes C, D, G, H, K, L were lysed after a 4-h chase in medium containing excess unlabelled methionine.

METHODS. *a*, 0.174, LBL 721, T2 and T1 cells were resuspended at 5×10^6 ml in medium (RPMI 1640/10% FCS) alone or containing peptides at 2×10^{-4} M for 5 h at 37 °C. They were then washed in 2 ml warm methionine-free medium and resuspended at 2×10^7 ml $^{-1}$ in methionine-free medium containing peptides at the same concentration for a further hour. [^{35}S]methionine (120 μCi) was then added and the mixture incubated for 15 min at 37 °C. Labelled cells were washed in ice-cold PBS and resuspended in 0.5 ml lysis buffer as described⁵. Lysates were precleared with *Staphylococcus a* organisms overnight at 4 °C and immunoprecipitated⁵ with purified MA2.1 antibody at a final concentration of $10 \mu\text{g ml}^{-1}$. Reduced immunoprecipitates were examined by 12% SDS-PAGE which was fixed, stained, treated with Amplify (Amersham), dried and exposed to pre-flashed X-ray film. *b*, T2 cells were resuspended at 5×10^6 ml $^{-1}$ in RPMI/10% FCS containing peptides at 2×10^{-4} M for 5 h. Aliquots of 2×10^7 T2 cells were resuspended in 1 ml methionine-free medium as above. After incubation for 1 h, 250 μCi [^{35}S]methionine was added for 15 min. The cells were then diluted to 5×10^6 ml $^{-1}$ by addition of 3 ml chase medium (RPMI, 10% FCS, 4 mM unlabelled methionine, 10 mM HEPES) containing peptides at the same concentration. One half of the cells were collected immediately, the remainder incubated for 4 h at 37 °C. Samples were washed, lysed in 1 ml lysis buffer and precleared as before. Immunoprecipitation was with purified MA2.1 antibody at a final concentration of $10 \mu\text{g ml}^{-1}$ (ref. 5). The endo H treatment has been described¹⁷.

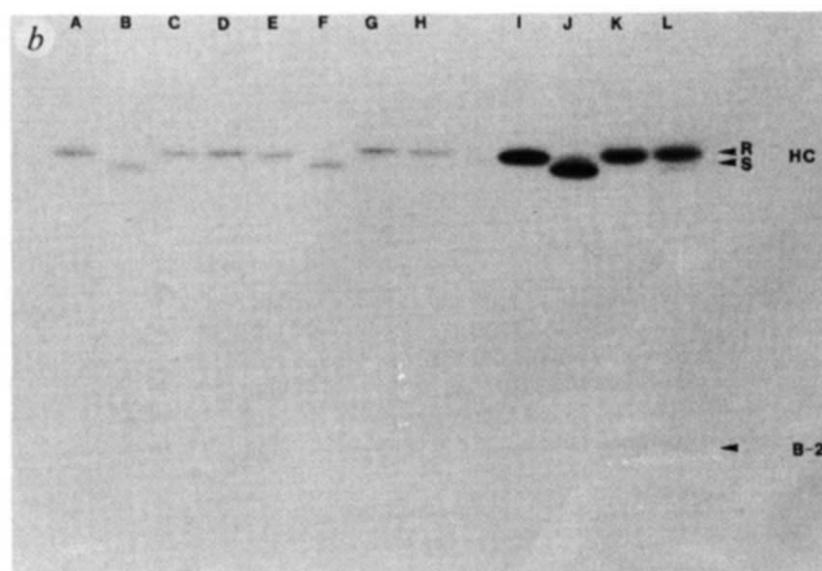


TABLE 1 Antigen presentation by T2 cells

Targets	Influenza-specific CTL restricted to		
	Kb	Aw68.1	B27
	% Specific lysis		
X31 + T2-Kb	7	ND	ND
X31 + C1R-Kb	50	ND	ND
H17 + T2-Aw68.1	ND	6	ND
NP89-101 + T2-Aw68.1	ND	67	ND
H17 + C1R-Aw68.1	ND	49	ND
NP89-101 + C1R-Aw68.1	ND	66	ND
X31 + T2-B27	ND	ND	2
NP383-390 + T2-B27	ND	ND	40
X31 + C1R-B27	ND	ND	85
NP383-390 + C1R-B27	ND	ND	98

The defect in antigen presentation in T2 is not limited to presentation of antigens with A2. T2 cells, transfected with class I genes coding for H-2 Kb, Aw68.1 or B27 were tested for presentation of viral antigens to influenza-specific CTL after virus infection or exposure to peptides. T2 cells and the control cell line C1R (an A, B null mutant LBL cell line⁹) were transfected with Kb, Aw68.1 and B27 as described⁹. Kb-restricted cytotoxic T lymphocytes (CTL) were generated from C57BL mice primed by intranasal infection with X31 influenza virus as described¹², and restimulated *in vitro* with C1R transfected with Kb and infected with X31 virus. The epitope recognized by these CTL has not been defined. Aw68.1 restricted CTL (V.C. and A.T., manuscript in preparation) stimulated by E61-13-H17 influenza^{13,14}, and B27 restricted CTL stimulated by X31¹⁵, recognize different epitopes mapped on the influenza nucleoprotein (NP) corresponding to the synthetic peptide indicated. The K to T ratio used was 10:1, and target cells were treated for 1 h with 5×10^{-5} M peptides. The per cent specific lysis was calculated as described¹⁶. ND, not determined.

soluble variant), and that this effect was peptide specific. The effect was blocked by treatment of the cells with brefeldin A (data not shown), implying that it depended on exit of class I molecules from the ER.

The effect of peptide on the conformation and transport of A2 molecules is shown in Fig. 3. Exposure to either peptide M58-68 or K62 (but not the control) increased the amount of A2 detected with MA2.1 (that binds folded A2 molecules associated with β -2microglobulin¹¹) in both 0.174 and T2 cells (Fig. 3a). However, total A2 detected with a serum raised to the cytoplasmic domain, was not altered by treatment with peptides (data not shown). Figure 3b shows that the peptides induced this conformational change on A2 heavy chains that are endo H-sensitive (lanes A, B and E, F), but which become resistant over a 4 h chase period (lanes C, D and G, H). This implies that peptide acts on A2 heavy chains before their entry into the Golgi apparatus.

These findings show that the phenotype of 0.174 and T2 cells is closely related to that of the murine mutant RMA-S⁵. The phenotype of these cells may be explained by the lack of available peptides during assembly of class I molecules in the endoplasmic reticulum. This might be due to loss of transport of peptides from the cytosol⁵, or possibly secondary to a defect in a peptide loading mechanism in the ER. Our results extend previous observations, and localize the mutation responsible for the loss of antigen presentation to the major histocompatibility complex on human chromosome 6. □

Received 8 February; accepted 27 March 1990.

- DeMars, R., Chang, C. C., Shaw, S., Reitnauer, P. J. & Sondel, P. M. *Hum. Immun.* **11**, 77-97 (1984).
- DeMars, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8183-8183 (1985).
- Shimizu, Y. et al. *Mol. cell. Biol.* **6**, 1074-1087 (1986).
- Salter, R. D. & Cresswell, P. *EMBO J.* **5**, 943-949 (1986).
- Townsend, A., Ohlen, C., Bastin, J., Foster, L. & Karre, K. *Nature* **340**, 443-448 (1989).
- Gotch, F., Rothbard, J., Howland, K., Townsend, A. & McMichael, A. *Nature* **326**, 881-882 (1987).
- Salter, R. D., Howell, D. N. & Cresswell, P. *Immunogenetics* **21**, 235-235 (1985).
- Whitehead, A. S., Colten, H. R., Chang, C. C. & DeMars, R. *J. Immun.* **134**, 641-641 (1985).
- Alexander, J., Payne, J. A., Murray, R., Frelinger, J. & Cresswell, P. *Immunogenetics* **29**, 380-388 (1989).
- Alexander, J., Payne, A. J., Shigekawa, B., Frelinger, J. & Cresswell, P. *Immunogenetics* **31**, 169-178 (1990).

- Ways, J. P., Rothbard, J. B. & Parham, P. *J. Immun.* **137**, 217-222 (1986).
- Townsend, A. R. M., Taylor, P. M., Mellor, A. L. & Askonas, B. A. *Immunogenetics* **17**, 283-294 (1983).
- Baez, M., Palese, P. & Kilbourne, E. D. *J. infect. Dis.* **141**, 362-365 (1980).
- Townsend, A. R. M. & Skehel, J. J. *J. exp. Med.* **160**, 552-563 (1984).
- Huet, S. et al. *Int. Immun.* (in the press).
- Townsend, A. R. M. et al. *Cell* **44**, 959-968 (1986).
- Owen, M. J., Kissnerghis, A. M. & Lodish, H. F. *J. biol. Chem.* **255**, 9678-9684 (1980).

ACKNOWLEDGEMENTS. We thank Dr R. DeMars for supplying the cell lines 0.174 and LBL721, Nigel Rust and the Nuffield Department of Surgery for analysing samples on the Ortho Cytofluorograf, and Diana Hayward for technical help.

Circular DNA is a product of the immunoglobulin class switch rearrangement

Uta von Schwedler, Hans-Martin Jäck & Matthias Wabl

Laboratory of Radiobiology and Environmental Health, and Department of Microbiology and Immunology, University of California, San Francisco, California 94143-0414, USA

THE class of immunoglobulin is defined by the constant region of its heavy chain. When a B lymphocyte switches the class of heavy chain it produces, the constant region of μ -type heavy chain is replaced (reviewed in refs 1 and 2, and see ref. 3); this occurs through a DNA rearrangement that brings the gene segment encoding the new constant region close to the VDJ segment encoding the variable region. The pre-B-cell line 18-81, which switches from heavy chain μ to γ 2b production in culture^{4,5}, occasionally abnormally rearranges the heavy chain locus so that DNA sequences between the switch regions of μ and γ 2b are inverted⁶. Because looping-out is an intermediate step in generating an inversion, the switch rearrangement could occur by looping-out and deletion. Provided that recombination is reciprocal, this would produce a circle of DNA. Indeed, circular DNA molecules have been isolated as products of rearrangement among gene segments encoding the variable regions of the T-cell receptor⁷⁻⁹ and of the immunoglobulin heavy chain^{10,11} and light chain^{11,12}. But whereas the breakpoints for the variable region rearrangement are precisely defined, the breakpoints for any given heavy chain class switch are scattered over a length of >6 kilobases, including both switch regions. We have now isolated circular DNA containing the sequences deleted by class-switching, thereby showing that the immunoglobulin heavy chain class switch occurs through looping-out and deletion.

A product of looping-out and deletion is a circle containing the constant region of μ -type heavy chain ($C\mu$) and the 3' part of the μ switch region joined to the 5' part of the switch (S) segment for the class to which the cell has switched (Fig. 1). These switch circles should be between 65 and 200 kilobases (kb) long, depending on the new class and on the exact location of the breakpoints.

As a likely source of switch circles, we used lipopolysaccharide-stimulated spleen cells, many of which had switched from μ to other heavy (H) chains¹³⁻¹⁵. After 3 days in culture, ~7% of the cells had switched to γ 3, 8% to γ 1, 3% to γ 2b, and 7% to γ 2a.

We purified circular DNA by the alkaline lysis method for mammalian cells^{16,17}. Lysis of the cells under highly basic conditions results in denaturation of linear (chromosomal) DNA, whereas the strands of covalently closed circular DNA remain intertwined. Extensive shearing of the cell lysate promotes efficient strand separation of the chromosomal DNA. After neutralization and phenol extraction at high salt concentration, the single-stranded DNA that arose from denaturation of the chromosomal DNA separates at the interphase. Covalently closed circular DNA is recovered from the aqueous phase,

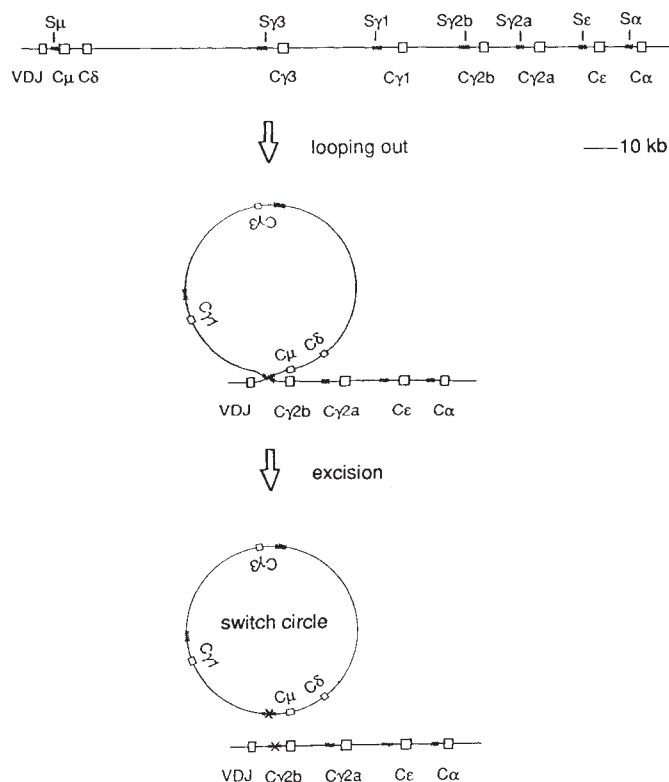


FIG. 1 Schematic representation of the H-chain locus showing organization of the constant gene segments, looping-out and deletion, and the resulting recombination products. In the mouse, the early immune response is dominated by IgM, which contains the H chain μ , and the later immune response is dominated by IgG3, IgG1, IgG2b, IgG2a, IgA and (rarely) IgE, which contain the H chains γ 3, γ 1, γ 2b, γ 2a, α and ϵ , respectively. Except for the constant region of δ -type H chain (C δ), each is preceded by a switch region of several kb within which the rearrangement breakpoint is usually found. X, switch recombination (break) point.

treated with RNase, and phenol-extracted again. Unlike caesium chloride density gradient centrifugation, this method allows isolation of large circles, including non-supercoiled ones, and takes only a few hours.

On Southern blots of circular DNA that had been digested with various restriction enzymes, we identified fragments that must be present in the switch circles. These fragments could not have been isolated from chromosomal DNA. Figure 2a shows

a restriction map of the joining segments of the H locus (JH) and C μ locus, which is needed to interpret the Southern blots (Fig. 2b and c). Figure 2b demonstrates the presence of C μ in the switch circle preparation. When hybridized with a C μ probe, the HindIII digest gave a 2.2-kb band and a 1.2-kb band in the circular DNA (lane 4), just as in the germ-line configuration (kidney) DNA (lane 3). Although at least 10 times more kidney DNA was applied to the gel, the bands of the circle preparation

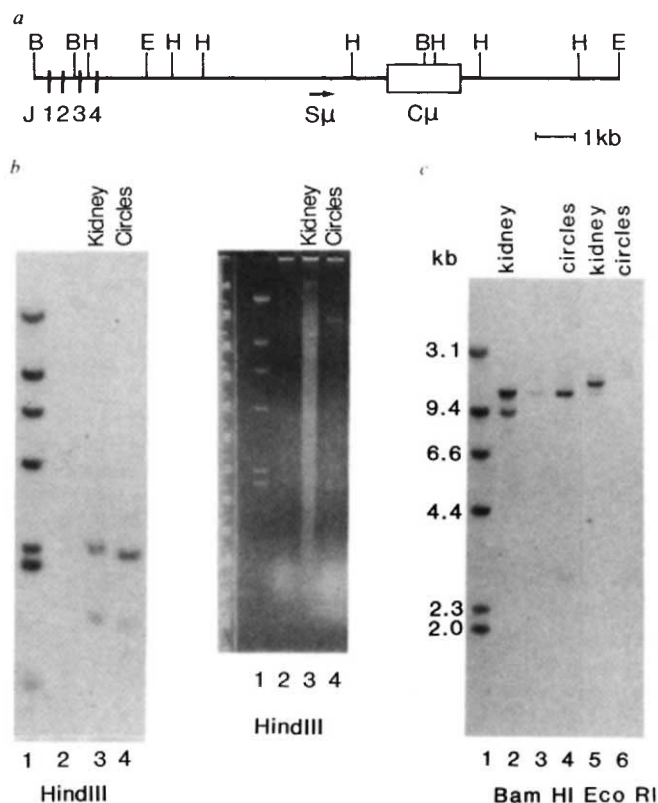


FIG. 2 Identification of the switch circles. *a*, Restriction enzyme map of embryonic BALB/c mouse DNA that includes the JH and C μ gene segments. B, BamHI; E, EcoRI; H, HindIII. *b*, Left: Southern blot analysis of kidney and circular DNA, digested with HindIII. The hybridization probe was C μ , a μ -gene complementary DNA fragment covering the first three exons. (The kidney DNA moves a little slower in the gel, a common occurrence in horizontal gels when excess DNA is loaded.) Lane 1, size marker. Right: ethidium bromide staining of the gel shown on the left. *c*, Southern blot analysis of kidney and circular DNA digested with BamHI lanes 2, 4 and EcoRI lanes 5, 6. The hybridization probe was C μ . Lane 1, size marker. Spleen cells of C57BL/6 mice were grown in RPMI medium containing 50 μ g ml⁻¹ lipopolysaccharide at a density of 2.5×10^5 cells ml⁻¹. On day 3, production of immunoglobulin isotypes was checked by membrane immunofluorescence. Circular DNA from 1.6×10^9 cells was isolated and dissolved in buffer (800 μ l). This solution (50 μ l, or the equivalent of 1×10^8 cells) was digested with restriction enzyme and loaded onto an agarose gel for Southern blot analysis. Chromosomal DNA (10 μ g, or the equivalent of 1×10^6 kidney cells) was loaded. Because there are two C μ copies per genome, this amount of DNA represents 2×10^6 copies of C μ . Circular and chromosomal DNA gave C μ bands of equal intensity. Thus, the circular DNA isolated from 1×10^8 spleen cells contains about 2×10^6 copies of C μ , or one switch circle per 50 cells. As only 30% of the spleen cells in culture had switched in isotype, this means that we could isolate a switch circle from 1 of 15 switched cells.

were as strong as those in the kidney DNA. Therefore, the $C\mu$ sequences in the circles could not all be due to chromosomal contamination unless preparation enriched for $C\mu$ sequences, an unlikely possibility. On the Southern blot shown in Fig. 2c, the *Bam*HI digest of kidney DNA (lane 2) gave a 9.4-kb band and an 11.2-kb band. The 9.4-kb band includes the sequences of *JH3* and *JH4*, as well as part of $C\mu$, whereas the 11.2-kb band covers the part of $C\mu$ 3' to the *Bam*HI site in $C\mu$. The circular DNA prepared from lipopolysaccharide-stimulated spleen cells (lane 4) shows the 11.2-kb band, which again confirms the presence of $C\mu$ in the switch circles. The lack of the 9.4-kb fragment in the *Bam*HI-digested circular DNA is also predicted for switch circles because the germ-line sequences 5' to $S\mu$ are replaced by the sequences 5' to the switch region of the other classes (Fig. 1). Because there are six other classes to which the μ -gene can switch, and because the breakpoints will be scattered over the switch regions and sometimes beyond, we expect the 9.4-kb band to be replaced with a smear. Indeed, a smear is present in lane 4, above the 11.2-kb band. If the 11.2-kb band were of chromosomal origin, the 9.4-kb band should also be present, albeit with half the intensity of the germ-line band in lane 2. (We assume that half of the spleen cells have rearranged their *VD* segment to *JH1* or *JH2*, and half to *JH3* or *JH4*. Rearrangements to *JH1* or *JH2* do not alter the *Bam*HI site 3' to *JH2*, thus preserving the 9.4-kb fragment; rearrangements to *JH3* or *JH4* delete the *Bam*HI site 3' to *JH2*.) Further evidence that the circular DNA is not contaminated with chromosomal DNA is given by the *Eco*RI digest in Fig. 2c. The 13.5-kb germ-line fragment (lane 5) is absent in the circular DNA (lane 6). The chromosomal *Eco*RI fragment containing $C\mu$ should be the same, whether derived from germ-line or rearranged DNA. Instead of the 13.5-kb band, a smear of bands larger than 13.5 kb is present in the circular DNA (lane 6). This is expected for the same reasons as those given for the *Bam*HI digest. Southern blot analyses of DNA digested with *Bgl*II and *Xba*I led to the same conclusions.

From the Southern blot analyses we estimate that we isolated switch circles from 1 of 15 switched cells (for calculation see Fig. 2). If the average clone size of switched cells is four to eight (two to three generations), then we recovered 0.25–0.5 circles per class switch event. In view of the losses to be expected from degradation *in vivo* and during isolation, this is a high yield. If the circles do not replicate, the yield is consistent with a circle arising from every H class-switch rearrangement.

We think that the circles generated by a switch to $\gamma 2b$ do not replicate as episomes. In clones of a pre-B-cell line in which both alleles have switched to $\gamma 2b$ (ref. 6), no $C\mu$ sequences

were detected by the polymerase chain reaction. If the circles replicated at the rate of the chromosomes, most of the $\gamma 2b$ -producing cells should contain at least one $C\mu$ segment.

We constructed a partial library of the *Eco*RI-digested switch circles in λ EMBL4 phage. On screening about 10^6 phages with a $C\mu$ probe we obtained 18 clones. We are aware that this procedure selects for a certain size-range of insert and probably excludes many fragments that contain breakpoints. As reasoned above, the breakpoints must be present on *Eco*RI fragments containing $C\mu$. The restriction maps of the DNA of the six different phage inserts analysed indicate that none was derived from chromosomal DNA (Fig. 3). Of five λ -inserts, we subcloned those fragments where we suspected the breakpoints to be, and sequenced parts of them. Four fragments had *Sy2b* joined to $S\mu$ sequences in a 5'–3' orientation (reciprocal to the 5'*S* μ –3'*Sy2b* configuration of the $\gamma 2b$ gene on the chromosome), and one fragment had *Sy3* joined to $S\mu$ (Fig. 4). Therefore, as predicted for the looping-out and deletion mechanism, the deleted DNA must have joined its ends to form a circle. The 5' sequence of the *Sy2b* clones is 94% concordant with a 210-base pair (bp) segment of a published part of the *Sy2b* sequence MUSIGCD23 (GenBank accession number J00456; ref. 18). The 3' sequence is 97% concordant with a 270-bp segment of the published $S\mu$ sequence MUSIGCD09 (GenBank accession no. J00442; ref. 19). Also, 144 bp of the 5' part of the fifth fragment are 98% concordant with the *Sy3* sequences MUSIGHAL (GenBank accession no. J00483; ref. 18) and MUSIGHANA (GenBank accession no. M12182; W. Dunnick, unpublished observations), whereas its 3' sequence is part of $S\mu$. We do not know whether the discrepancies reflect polymorphism (the GenBank sequences are derived from the BALB/c mouse strain, our circle sequences from the C57BL/6 strain) and/or sequencing errors. Although more spleen cells had switched to other γ -subclasses than to $\gamma 2b$, four out of five switch-recombination sequences joined to $S\mu$ were from *Sy2b*. This is probably due to our selection of *Eco*RI fragments for the library.

In principle, chimaeric switch regions, that is, 5'*Sy*–3' $S\mu$, could also be generated by recombination between sister chromatids or homologues. In these cases, however, the chimaeric switch regions would be located on chromosomal DNA, not on circles. Nevertheless, as the following calculation demonstrates, our results do not depend on any *a priori* estimate of the purity of our preparation of circular DNA. Immediately after the putative unequal crossing-over event, half, at most, of the $S\mu$ regions would be chimaeric. This fact is not altered by growth of the cells after the switch. Because no more than 30% of our cells switched, no more than 15% of the switch regions

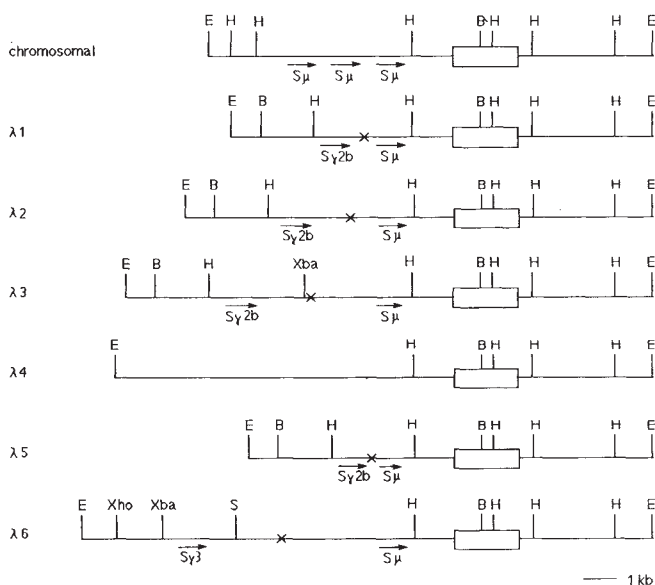


FIG. 3 Restriction maps of the chromosomal μ -locus and of 6 of the 18 λ -clones containing inserts with the breakpoints from switch circles. E, *Eco*RI; H, *Hind*III; B, *Bam*HI; S, *Sph*I; Xho, *Xho*I; Xba, *Xba*I. X, putative breakpoints. The arrows at the *Sy2b*, *Sy3* and $S\mu$ sequences show their original orientation at the H locus.

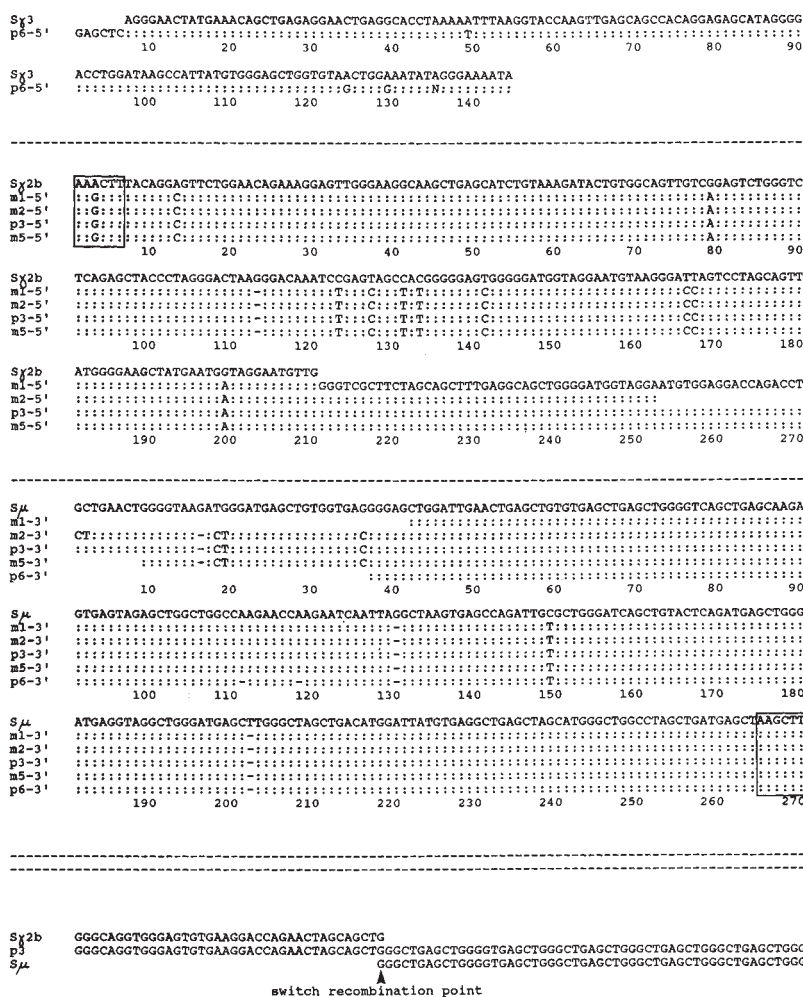


FIG. 4 *a*, Sequences of the 5' and 3' ends of five switch-circle fragments that contain the switch recombination points. Clones m1, m2 and m5 contain *Hind*III fragments of the λ -phages λ 1, λ 2 and λ 5, respectively, subcloned into M13mp19; clone p3 contains a *Hind*III fragment of λ 3 in pGem3, and pGem clone p6 contains the *Eco*RI-*Hind*III fragment of λ 6. The sequences were aligned with a $S\gamma$ 3 switch-region sequence (MUSIGHAL) or with a $S\gamma$ 2b sequence (MUSIGCD23) and with part of the $S\mu$ sequence (MUSIGCD09). $\cdot$, base identity; —, a gap. Nucleotides 211–270 of the $S\gamma$ 2b sequence were not present in the GenBank. The *Hind*III restriction sites AAGCTT, that were used for subcloning, are boxed. *b*, Sequence of the switch recombination (break) point of clone p3. The 5' part is identical to the $S\gamma$ 2b sequence MUSIGCD24 (GenBank accession no. J00457; ref. 18), the 3' part represents the $S\mu$ consensus sequence, GAGCTGGGG/CT.

would be chimaeric. Even if the putative circular DNA were entirely linear the probability of our isolating six chimaeric out of a total of six switch regions would be 0.15^6 , or ~ 0.00001 . So it is reasonable to conclude that the chimaeric switch regions that we isolated are derived from circular DNA.

The existence of switch circles is evidence that looping-out and deletion occur during the heavy chain class-switch in mitogen-stimulated B lymphocytes. By itself, this does not exclude the possibility that two other proposed mechanisms might also achieve the same result: unequal exchange between sister chromatids^{20,21} and between homologues. Illegitimate recombination between the switch regions and regions on other chromosomes must occur, albeit at a low frequency; for example, oncogenes are translocated to the immunoglobulin locus²². Class switching of a randomly integrated μ -transgene has also been reported²³, implying interchromosomal recombination, or *trans*-splicing of RNA. But we have excluded the possibility that either the exchange between sister chromatids²⁴ or the exchange between homologues^{6,24} makes any substantial contribution to the heavy chain class-switch in a pre-B-cell line. Also unequal exchange between homologues has been excluded by the fact that the V and C region allotypes of a given immunoglobulin are both encoded by one homologue in the rabbit^{25–28} and in the mouse²⁹. We also note that, because in the mouse the order on the chromosome is centromere, C, V (ref. 30), recombination between homologues would create switched cells, half of which would contain two $C\mu$ segments on the silent homologue; the other half of switched cells would be μ - γ double producers with identical V regions on each isotype. These types of cells have not been demonstrated, which, in view of the many hybridomas analysed, is another strong argument against this

mechanism making a significant contribution to the switch process.

As in the deletional joining of the V, D, and J segments of the T-cell receptor and immunoglobulin genes, the switch recombination includes a reciprocal ligation resulting in circular DNA. It has been proposed that circles are less recombinogenic with respect to chromosomal integration than linear molecules and thus minimize potential mutagenic damage³¹. □

Received 9 February; accepted 26 April 1990.

- Shimizu, A. & Honjo, T. *Cell* **36**, 801–803 (1984).
- Marcu, K. B. *Cell* **29**, 719–721 (1982).
- Shimizu, A., Takahashi, N., Yaoita, Y. & Honjo, T. *Cell* **28**, 499–506 (1982).
- Burrows, P. D., Beck, G. B. & Wabl, M. R. *Proc. natn. Acad. Sci. U.S.A.* **78**, 564–568 (1981).
- Burrows, P. D., Beck-Engeser, G. B. & Wabl, M. R. *Nature* **306**, 243–246 (1983).
- Jäck, H.-M., McDowell, M., Steinberg, C. M. & Wabl, M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1581–1585 (1988).
- Fujimoto, S. & Yamagishi, H. *Nature* **327**, 242–243 (1987).
- Okazaki, K., Davis, D. D. & Sakano, H. *Cell* **49**, 477–485 (1987).
- Okazaki, K. & Sakano, H. *EMBO J.* **7**, 1669–1674 (1988).
- Abe, M. & Shikui, H. *Nucleic Acids Res.* **17**, 163–170 (1989).
- Toda, M., Hirma, T., Takeshita, S. & Yamagishi, H. *Immun. Lett.* **21**, 311–316 (1989).
- McCormack, W. T. *et al. Cell* **56**, 785–791 (1989).
- Kearney, J. F. & Lawton, A. R. *J. Immun.* **115**, 671–681 (1975).
- Pernis, B., Fornì, L. & Luzzati, A. L. *Cold Spring Harb. Symp. quant. Biol.* **41**, 175–183 (1976).
- Wabl, M. R., Fornì, L. & Loo, F. *Science* **199**, 1078–1080 (1978).
- Griffin, B. E., Björck, E., Bjursell, G. & Lindahl, T. *J. Virol.* **40**, 11–19 (1981).
- Carroll, S. M. *et al. Molec. cell. Biol.* **7**, 1740–1750 (1987).
- Kataoka, T., Miyata, T. & Honjo, T. *Cell* **23**, 357–368 (1981).
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676–683 (1980).
- Obata, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **78**, 2437–2441 (1981).
- Tilley, S. A. & Birshstein, B. K. *J. exp. Med.* **162**, 675–694 (1985).
- Altick, E. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 6333–6337 (1989).
- Durdik, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 2346–2350 (1989).
- Wabl, M., Meyer, J., Beck-Engeser, G., Terkhoff, M. & Burrows, P. D. *Nature* **313**, 687–689 (1985).
- Kindt, T. J., Mandy, W. J. & Todd, C. W. *Biochemistry* **9**, 2028–2032 (1970).
- Landucci Tosi, S., Mage, R. G. & Dubiski, S. *J. Immun.* **104**, 641–647 (1970).
- Knight, K. L. & Hanly, W. C. *Contemp. Top. molec. Immun.* **4**, 55–58 (1975).

ACKNOWLEDGEMENTS. We thank Dr Susan Carroll for help with the alkaline lysis method, Dr Philip Tucker for DNA probes and for unpublished immunoglobulin gene sequences, Dr Charles Steinberg for discussions, Pam Winegar for typing the manuscript, and Mary McKenney for editing it. This work was supported by the NIH and the Office of Health and Environmental research, US Department of Energy. U.V.S. is a fellow of the Boehringer-Ingelheim Fonds, West Germany.

Sy-Shi Wang & Virginia A. Zakian*

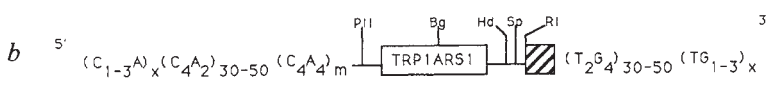
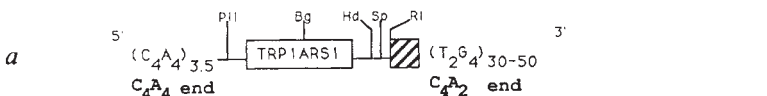
Fred Hutchinson Cancer Research Center, 1124 Columbia Street, Seattle, Washington 98104, USA

* To whom correspondence should be addressed

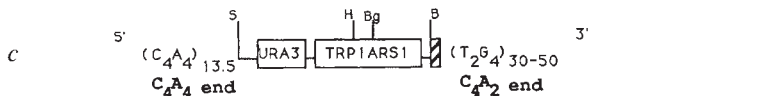
DNA termini from *Tetrahymena* and *Oxytricha*, which bear C₄A₂ and C₄A₄ repeats respectively, can support telomere formation in *Saccharomyces cerevisiae* by serving as substrates for the addition of yeast telomeric C₁–₃A repeats. Previously, we showed that linear plasmids with 108 base pairs of C₄A₄ DNA (YLp108CA) efficiently acquired telomeres, whereas plasmids containing 28–64 base pairs of C₄A₄ DNA also promoted telomere formation, but with reduced efficiency¹. Although many of the C₄A₄ termini on these plasmids underwent recombination with a C₄A₂ terminus, the mechanism of telomere–telomere recombination was not established^{1–3}. We now report the sequence of the C₄A₄ ends from the linear plasmids. The results provide strong evidence for a novel recombination process involving a gene conversion event that requires little homology, occurs at or near the boundary of telomeric and non-

A

Plasmid YLp28CA



Plasmid YLp108CA



B

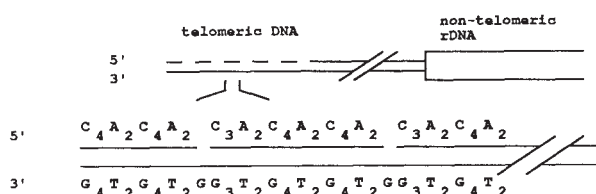


Figure 1a shows the structure of linear plasmids before and after introduction into yeast (before transformation, there are no detectable C_4A_2 or C_4A_4 sequences in yeast¹). We were able to distinguish yeast transformants where recombination at the C_4A_4 end had occurred, and those that did not undergo recombination, by Southern hybridization (Table 1). C_4A_4 termini were cloned and the junction between telomeric and non-telomeric DNA was determined. In each recombinant, more than 180 base pairs (bp) of C_4A_2 sequences were found at the C_4A_4 end (Fig. 2c); the polarity of the transferred C_4A_2 sequences was the same as that of the C_4A_4 repeats (Fig. 2b); and there was no detectable decrease in the amount of C_4A_2 DNA at the C_4A_2 end (data not shown).

These results have important implications for the mechanism of telomere-telomere recombination. First, the fact that a long stretch of C_4A_2 DNA is transferred, an amount indistinguishable from the amount of uninterrupted C_4A_2 DNA at the C_4A_2 end (Fig. 1b), argues that initiation (or resolution) of recombination at the C_4A_2 end must have occurred near or at the junction between the C_4A_2 repeats and non-telomeric DNA (Fig. 3). These data rule out models^{1,2} in which recombination results from the transfer of repeats between nicks in the C_4A_2 strand (Fig. 1b) which predict that only short tracts of C_4A_2 should be transferred.

Second, these results rule out the possibility that transfer occurred by the resolution reaction. In this model^{3,4}, a C_4A_4 and a C_4A_2 end ligate together followed by breakage within the C_4A_2 sequences. Contrary to the observed results (Fig. 2b), the resolution reaction predicts that the orientation of the transferred C_4A_2 repeats would be inverted relative to that of the C_4A_4 end repeats^{1,3}.

Third, the fact that there is no detectable decrease in the number of C_4A_2 repeats at the C_4A_2 end indicates that recombination usually occurred by non-reciprocal recombination.

Intramolecular reciprocal recombination would result in loss of >180 bp of C₄A₂ DNA from the C₄A₂ end, a result that is

FIG. 1 A, Structures of plasmids Ylp28CA and Ylp108CA before (a, c) and after (b, d) introduction into yeast. One strand is represented in the 5' to 3' orientation (left to right). Construction of the linear plasmids has been described¹. Plasmids Ylp28CA and Ylp108CA bearing 28 and 108 bp of C₄A₄ repeats, respectively, at the 'C₄A₄ end' and ~300 bp of C₄A₂ repeats at the 'C₄A₂ end' were introduced by transformation²¹ into yeast. The structure of plasmid DNA in individual transformants carrying linear plasmids was determined by restriction enzyme mapping and Maxam and Gilbert DNA sequencing²². As expected, all linear plasmids acquired yeast C₁-₃A repeats at each end. Many (90%) plasmids had also undergone recombination with other termini as demonstrated by the transfer of tracts of C₄A₂ repeats to the C₄A₄ end. Sometimes (~50% of sequenced plasmids), the remaining C₄A₄ end length (C₄A₄)_m [or (C₄A₄)_n] was less than 28 bp (or 108 bp) (see Table 1). The length of (C₁-₃A)_x or (TG₁₋₃)_x is ~300 bp. The hatched box indicates the C₄A₂-adjacent unique sequence from *Tetrahymena* rDNA. Abbreviations: Bg, *Bgl*II; B, *Bam*HI; H, *Hind*III; Pli, *Pvu*II; RI, *Eco*RI; S, *Sal*I; Sp, *Sph*I. B, Structure of the terminus of *Tetrahymena* extrachromosomal rDNA molecules. These termini bear an average of 300 bp of C₄A₂ sequence. Specific single-base gaps occur every 12-18 nucleotides in the most terminal 100 bp of the molecule²³. The G-rich strand is extended as a 12 to 16-nucleotide single-strand tail on at least some termini²⁴.

TABLE 1 Sequences at C₄A₄-end junctions

Yeast transformant	No. bp C ₄ A ₄ retained	No. bp C ₄ A ₂ transferred	Approximate bp C ₁₋₃ A	Sequence at the junction
Ylp108CA (without transfer)				
1	108	0	300	ACACAT ₂ C ₃ C ₄ A ₄
2	84	0	300	CACACA C ₄ A ₄
3	76	0	400	CACACA C ₄ A ₄
4	36	0	300	CCCACA C ₄ A ₄
Ylp40CA				
1	36	>180	300	C ₄ A ₂ C ₄ A ₄
2	36	>180	300	C ₄ A ₂ C ₄ A ₄
3	12	>180	300	C ₄ A ₂ C ₄ A ₄
Ylp28CA				
1	20	>240	380	C ₄ A ₂ C ₄ A ₄
2	12	>240	270	C ₄ A ₂ C ₄ A ₄
3	28	>240	270	C ₄ A ₂ C ₄ A ₄
4	28	>240	270	C ₄ A ₂ C ₄ A ₄

Linear plasmids with 108, 40, or 28 bp of cloned C₄A₄ DNA at the C₄A₄ end were introduced and propagated in yeast (see Fig. 1). The C₄A₄-end telomeres of plasmids Ylp108CA and Ylp28CA were cloned from various yeast transformants and sequenced as described (Fig. 2). The C₄A₄-end telomeres of plasmid Ylp40CA were cloned in the same way as for Ylp28CA. For each yeast transformant, two independent *Escherichia coli* telomere clones were sequenced. In each case, the amount of C₄A₄ DNA retained and the junction sequence between C₄A₄ and C₁₋₃A (or C₄A₂) were identical between the two clones. Plasmids that underwent recombination obtained over 180 bp of C₄A₂ DNA at the C₄A₄ end before the addition of C₁₋₃A. The approximate length of added C₁₋₃A sequence was estimated by agarose gel electrophoresis or in some cases determined by DNA sequencing. Sequences at the junction between the C₄A₄ DNA and the added C₁₋₃A or C₄A₂ DNA are shown in the 5' to 3' orientation (left to right) running from the end towards the central portion of the linear plasmids.

never observed. Although 20% of the recombinants had C₄A₂ and C₄A₄ DNA on both ends, which is the product expected for reciprocal intramolecular recombination, even these molecules have no detectable loss of C₄A₂ DNA at the C₄A₂ end¹. We cannot rule out the possibility that some recombinants are the result of reciprocal recombination between two different plasmid molecules. But we never recovered the reciprocal product of inter-molecular recombination in the same cell despite the fact that as long as this product has ~28 bp of telomeric DNA remaining at the C₄A₂ end, it should be replication-competent. Moreover, since reciprocal recombination will result in a reduction in the amount of C₄A₄ DNA at the C₄A₄ end, inter-molecular reciprocal recombination cannot account for the two Ylp28CA transformants that retained the full 28 bp of C₄A₄ DNA after recombination. Recombination on these plasmids must have occurred by a non-reciprocal event.

When linear plasmid DNA is introduced into yeast, its ends are believed to be subject to exonucleolytic degradation. DNA sequencing demonstrates that even termini bearing telomeric repeats are subject to degradation: 5 of 7 of the sequenced C₄A₄ ends that had acquired C₄A₂ repeats (and 3 of 4 termini that had not acquired C₄A₂ repeats; see below) had fewer C₄A₄ repeats than the input plasmid (Table 1). Two termini that had undergone recombination retained only 12 bp of C₄A₄ DNA. Therefore, telomere-telomere recombination must require very little homology (no more than 28 bp and possibly as few as 12 bp). In contrast, for more conventional yeast sequences, 26 bp

of perfect homology supports very little gene conversion (10⁻¹⁰ events per generation)⁵.

We also examined the termini of plasmids that did not recombine. Ylp108CA did not undergo recombination in 4 out of 19 transformants¹. DNA sequencing of cloned C₄A₄ termini from these transformants revealed that no C₄A₂ repeats were transferred to the C₄A₄ end. These results confirm our earlier results⁶ that C₄A₄ repeats are themselves suitable substrates for C₁₋₃A addition in yeast. The junction between C₄A₄ DNA and C₁₋₃A DNA was determined for all four termini (Table 1). In one case, all of the C₄A₄ repeats on the input plasmid were retained and the C₁₋₃A repeats were added directly to the five bases of non-telomeric DNA that are the remnant of the restriction enzyme recognition site used to generate the C₄A₄ end (Table 1; Fig. 2a). This result demonstrates that telomeric DNA need not be at the end of a DNA molecule to be utilized for telomere formation.

The same conclusion was reached by others from restriction enzyme digestion of yeast linear plasmid DNA⁴. In this earlier study, the stretch of telomeric DNA (75 repeats of alternating CA sequence) was 66 bp from the end of the input linearized plasmid, and ~40% of the termini retained at least 35 bp of non-telomeric DNA after telomere formation. It has been argued that the addition of C₁₋₃A repeats to termini capped by non-telomeric DNA can not occur by recombination⁴. However, recombination reactions both *in vitro*⁷ and *in vivo*⁸ can bypass a heterology. It is therefore possible that telomere formation,

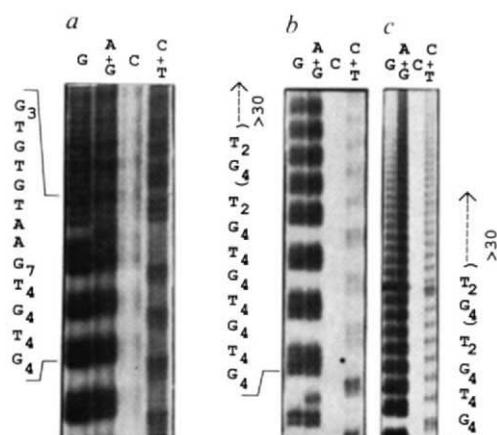


FIG. 2 DNA sequence of C₄A₄ termini. Sequences of C₄A₄ termini from plasmids derived from two independent yeast transformants are shown. a. Sequence of the junction between G₄T₄ repeats and the added G₁₋₃T repeats in a plasmid that did not undergo recombination, Ylp108CA-1. b. Sequence of the junction between G₄T₄ and G₄T₂ repeats in a plasmid that underwent recombination, Ylp28CA-3. c. Long tract of G₄T₂ repeats (over 30 copies) are detected at the C₄A₄ end in clone Ylp28CA-3.

METHODS. Telomere clones Ylp108CA-1 and Ylp28CA-3 were constructed as follows. Total DNA from yeast cells containing the linear plasmids Ylp108CA or Ylp28CA was digested with *Bgl*II or *Sph*I, respectively, treated with T4 DNA polymerase in the presence of dNTPs and self-ligated. Ligation mixtures were used to transform *E. coli* strain BSJ72 to ampicillin resistance. Telomere clones derived from Ylp108CA were further digested with *Hind*III and *Sal*I and the fragments containing the telomeric sequences were subcloned into *Hind*III and *Sal*I digested plasmid pVZ1²⁵. For each yeast transformant, telomere clones were isolated and sequenced from two independent *E. coli* transformants. In each case, the sequence at the junction between G₄T₄ repeats and the added G₁₋₃T or G₄T₂ repeats was found to be identical. Sequencing analysis was conducted by the method of Maxam and Gilbert²².

even on termini capped by a stretch of non-telomeric DNA, involves recombination, especially if the last few bases of the substrate are able to pair with $C_{1-3}A$ sequence. Alternatively, if telomerase⁹⁻¹² exists in *Saccharomyces* and if it can use a substrate in which the telomeric sequences are subterminal, it could mediate those telomere formation events that do not involve recombination.

Our results indicate that telomere formation in yeast is accompanied by telomere-telomere recombination. This recombination requires surprisingly little homology, perhaps because it is a site-specific event promoted by the ability of telomere DNA to assume non-B form structures involving triplex or quadruplex associations^{13,14} or G-G¹³⁻¹⁵ or C-C base pairing¹⁶. The junction between telomeric and unique sequence DNA, an apparent hot spot of recombination *in vivo*, also acts as a boundary for Klenow DNA polymerase and S1 nuclease *in vitro*^{17,18}. Although the data here and elsewhere⁴ do not rule out alternative pathways for telomere formation in yeast, telomere-telomere recombination provides an efficient mechanism for immediate rescue of DNA termini with very short stretches of telomere DNA. In contrast, other events that alter telomere length in yeast do so gradually such that many generations are needed to see a change

in telomere length (reviewed in ref. 19). A working model for telomere formation in yeast by recombination is shown in Fig. 8. This scheme is reminiscent of the bacteriophage T4 'copy choice' model^{8,20} that allows complete replication of the T4 genome by a process that requires recombination. As this recombination is non-reciprocal and can result in a net increase in telomere sequences it could contribute to telomere replication during the normal cell cycle.

If telomere-telomere recombination is important for telomere replication or maintenance, we also expect authentic telomeres to recombine. However, we did not detect transfer of C_4A_4 or C_4A_2 DNA to yeast chromosomes. These results can be explained if yeast telomeres are more likely to recombine with each other than with plasmid termini. This is possible as (1) chromosomal telomeres and internal stretches of $C_{1-3}A$ are more abundant than plasmid termini, and (2) chromosomal telomeres have much greater homology to each other than to C_4A_4 or C_4A_2 DNA. Alternatively, if telomere-telomere recombination is a salvage pathway that acts only on those telomeres with very short stretches of $C_{1-3}A$ DNA, its occurrence at natural chromosomes may be too rare to detect by Southern hybridization. □

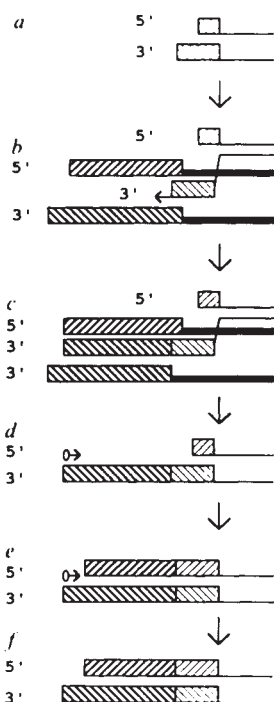


FIG. 3 Model for telomere formation by recombination. Different telomeric repeat sequences (thin-line hatched box) can promote telomere formation in yeast by serving as substrates for the addition of yeast $C_{1-3}A$ repeats (thick-line hatched box). In both cases, the C-rich strand is marked with 5' and the G-rich strand is marked with 3'. Only one terminus of either a linear plasmid or a chromosome is shown. *a*, After replication of the plasmid in yeast and removal of the RNA primer, a single-strand tail is left at the 3' end of the newly replicated strand (represented by the unpaired thin-line hatched box). *b*, The 3' OH end of the single-strand G-rich tail invades another telomere (donor, represented by thick-line hatched box). At the donor end, most (or all) recombination events are initiated (or resolved) at the junction between telomeric DNA (thick-line hatched box) and unique DNA (filled-in bar). *c*, The invading G-rich strand is extended by replication using the donor telomere as a template. *d*, After dissociation, the terminus carries an ~300 nucleotide long $G_{1-3}T$ single-strand tail and serves as a template for primase (O) and conventional DNA polymerase mediated replication of the complementary strand (e). *f*, Subsequent removal of the RNA primer would still leave a gap at the 5' end of the newly replicated strand, but no sequence information would be lost. (Alternatively, the extended G-rich strand could fold back on itself to form a terminal hairpin and provide the primer for replication of the C-rich strand¹⁵.)

Received 21 December 1989; accepted 16 March 1990

- Pluta, A. & Zakian, V. *Nature* **337**, 429-433 (1989).
- Walmsley, R., Chan, C., Tye, B.-K. & Petes, T. *Nature* **310**, 157-160 (1984).
- Szostak, J. W. *Nature* **337**, 303-304 (1989).
- Murray, A., Claus, T. & Szostak, J. *Molec. cell. Biol.* **8**, 4642-4650 (1988).
- Ahn, B.-Y., Dornfeld, K. J., Fagrellius, T. J. & Livingston, D. M. *Molec. cell. Biol.* **8**, 2442-2448 (1988).
- Pluta, A., Dani, G., Spear, B. & Zakian, V. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1475-1479 (1984).
- Bianchi, M. E. & Radding, C. M. *Cell* **35**, 511-520 (1983).
- Brunier, D., Michel, B. & Ehrlich, S. D. *Cell* **52**, 883-892 (1988).
- Greider, C. W. & Blackburn, E. H. *Cell* **43**, 405-413 (1985).
- Zahler, A. M. & Prescott, D. M. *Nucleic Acids Res.* **16**, 6953-6972 (1988).
- Shippen-Lentz, D. & Blackburn, E. H. *Molec. cell. Biol.* **9**, 2761-2764 (1989).
- Morin, G. *Cell* **59**, 521-529 (1989).
- Sen, D. & Gilbert, W. *Nature* **334**, 364-366 (1988).
- Williamson, J. R., Raghuraman, M. K. & Cech, T. R. *Cell* **59**, 871-880 (1989).
- Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I. Jr & Blackburn, E. H. *Cell* **51**, 899-908 (1987).
- Lyamichev, V. I. *et al. Nature* **339**, 634-637 (1989).
- Henderson, E. *et al. in Eukaryotic DNA Replication Vol. 6*, (eds Stillman, B. & Kelley, T.) 453-461 (Cold Spring Harbor Laboratory, New York, 1988).
- Budarf, M. & Blackburn, E. *Nucleic Acids Res.* **15**, 6273-6292 (1987).
- Zakian, V. A. *A. Rev. Genet.* **23**, 579-604 (1989).
- Mosig, G. A. *Rev. Genet.* **21**, 347-371 (1987).
- Dunn, B., Szauter, P., Pardue, M. L. & Szostak, J. W. *Cell* **39**, 191-201 (1984).
- Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).
- Blackburn, E. H. & Gall, J. *J. molec. Biol.* **120**, 33-53 (1978).
- Henderson, E. R. & Blackburn, E. H. *Molec. cell. Biol.* **9**, 345-348 (1989).
- Henikoff, S. & Eghtedarzadeh, M. K. *Genetics* **117**, 711-725 (1985).

ACKNOWLEDGEMENTS. We thank G. Smith, A. Taylor, A. MacAuley and the members of our laboratory for comments on the manuscript. This work is supported by the NIH (V.A.Z.) and the Joan Taylor Guild (S.-S.W.).

Telomeres shorten during ageing of human fibroblasts

Calvin B. Harley*, A. Bruce Futcher†
& Carol W. Greider†

* Department of Biochemistry, McMaster University,
1200 Main Street West, Hamilton, Ontario L8N 3Z5, Canada
† Cold Spring Harbor Laboratory, Cold Spring Harbor,
New York 11724, USA

THE terminus of a DNA helix has been called its Achilles' heel¹. Thus to prevent possible incomplete replication² and instability^{3,4} of the termini of linear DNA, eukaryotic chromosomes end in characteristic repetitive DNA sequences within specialized structures called telomeres⁵. In immortal cells, loss of telomeric DNA due to degradation or incomplete replication is apparently balanced by telomere elongation⁶⁻¹⁰, which may involve *de novo* synthesis of additional repeats by a novel DNA polymerase called telomerase¹¹⁻¹⁴. Such a polymerase has been recently detected in HeLa cells¹⁵. It has been proposed that the finite doubling capacity

TABLE 1 Effect of donor age on telomere length in human fibroblasts

Cell strain	Age		Mean telomere length kb $\pm$ s.d. (n)
	<i>in vivo</i> years	<i>in vitro</i> MPD (MPD max)	
HSC172	Fetal	18–28 (88)	8.6 $\pm$ 0.5 (3)
A30S	0	33 (58)	7.3 (1)
A38	24	31–33 (68)	6.9 $\pm$ 0.3 (2)
A35	70	19 (41)	6.7 (1)
F001	71	21–29 (40)	6.5 $\pm$ 0.4 (5)
F002	91	18–20 (45)	6.2 $\pm$ 0.1 (3)

Mean telomere length (the length of the terminal restriction fragment) was determined as described in Fig. 2a for fibroblast cell strains at the earliest available mean population doubling (MPD) in separate experiments. Strains were derived from female fetal lung (HSC172), female newborn skin (A30S), male forearm skin (A38, A35) or female abdominal skin (F001, F002). MPD at time of assay and senescence (MPD max) are indicated. The correlation between increasing donor age and decreasing telomere length is statistically significant ($P < 0.05$).

of normal mammalian cells is due to a loss of telomeric DNA and eventual deletion of essential sequences^{1,16,17}. In yeast, the *est1* mutation causes gradual loss of telomeric DNA and eventual cell death mimicking senescence in higher eukaryotic cells¹⁷. Here, we show that the amount and length of telomeric DNA in human fibroblasts does in fact decrease as a function of serial passage during ageing *in vitro* and possibly *in vivo*. It is not known whether this loss of DNA has a causal role in senescence.

The finite replicative capacity of human fibroblasts in culture¹⁸ is well established and has been extensively used as a model of cellular senescence¹⁹. To investigate the effect of cell division and age on telomeric DNA, we grew nontransformed human fibroblasts *in vitro* until terminal passage and regularly determined the size distribution of terminal restriction fragments (that is, telomeres) by Southern blot analysis (Fig. 1). Mean telomere length decreased about 2 kilobases (kb) with cumulative population doublings in five strains of fibroblasts studied

(Fig. 2a). Moreover, the total amount of telomeric DNA also decreased (Fig. 2b), which indicates true loss of DNA and not simply rearrangement of telomeres with respect to restriction enzyme cleavage sites. Loss of telomeric DNA did not result from general degradation or loss of repetitive DNA in preparations of DNA from old cells²⁰, as other repetitive, non-telomeric sequence elements analysed in these experiments were not altered in size or amount (Fig. 1b, d). The loss was also not a function of growth state: cells at similar passage levels collected during exponential growth or during quiescence showed no significant difference in telomere distribution (not shown).

As loss of telomeric DNA observed with ageing *in vitro* may occur generally as somatic cells divide *in vivo*, it is important to determine the effect of donor age and tissue type on telomere length. Telomere length in fibroblasts varies between individuals and may become shorter during ageing *in vivo* (Figs 1 and 2, and Table 1). The increased heterogeneity of telomere lengths in cells from old donors (Fig. 1c, g) presumably reflects their variable history *in vivo* as well as their increased age in cell doublings. A larger population of individuals and other tissues must, however, be studied with both repetitive and unique-sequence telomere probes to confirm these observations. In addition, comparative studies should attempt to correlate telomere loss with cellular life span in different species.

The degree of telomere shortening observed during ageing of cultured human fibroblasts could be biologically significant. The ends of human telomeres consist of repeats of the sequence TTAGGG which are probably functionally important²¹. Human sperm telomeres have about 9 kb of TTAGGG repeats, whereas somatic cell telomeres may have only 4 kb. In addition, there seem to be variant forms of TTAGGG (such as TTGGGG) present sub-terminally²¹. The rates of loss of length and amount of telomeric DNA we have observed (Fig. 2) are in rough agreement with the estimate of 4 kb of TTAGGG per 'young' somatic telomere. Thus, the loss of 2 kb of terminal sequences could represent loss of most of the functional telomeric sequences. Even if there is substantially more than 4 kb of functional sequence at telomeres from young cells, this shortening

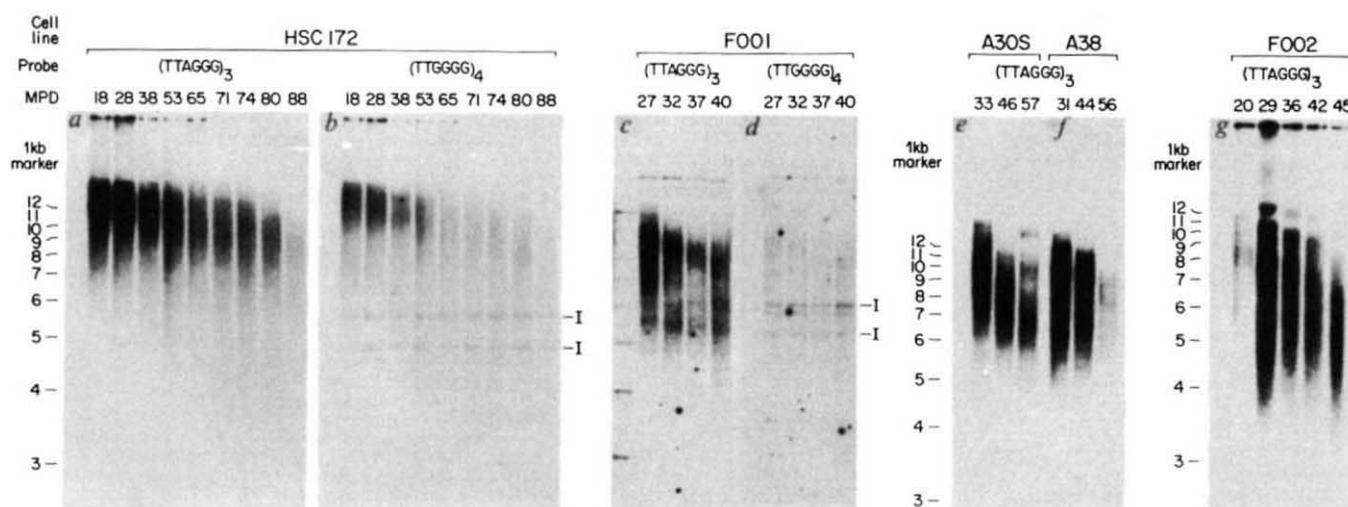


FIG. 1 Terminal restriction fragment (telomere) length in cultured human fibroblasts. Genomic DNA from fetal cell strain HSC172 (a, b), old adult F001 (c, d), F002 (g), newborn A30S (e) and young adult A38 (f) cell strains (see Table 1) was analysed at indicated mean population doublings (MPD) on Southern blots. Oligonucleotide probe (TTAGGG)₃ was used with high stringency washes to specifically detect telomeric sequences (a, c, e, f, g) and oligonucleotide (TTGGGG)₄ with low stringency washes to detect both telomeric sequences and internal repetitive sequences (I) (b, d). The size (kb) and position of markers are indicated.

METHODS. Cells were grown and population doublings counted as described³¹ and DNA isolated when cells reached confluence. DNA was

digested to completion with restriction enzymes *MspI* and *RsaI* and carefully quantified by a fluorometric method³². A portion (1–2 μ g) was loaded onto a 0.7% agarose gel and resolved by electrophoresis at 1 V cm⁻¹ for 36–72 h. For F002, only 1 μ g was loaded in the first lane (MPD 20), 4 μ g in the second lane (MPD 29) and 2 μ g in each of the other lanes. DNA was depurinated by soaking gels in 0.1 M sodium citrate (pH 3.0) for 2 h, transferred to Nytran (Schleicher and Schuell) and hybridized in 6 $\times$ SSC at 37 $^{\circ}$ C or 50 $^{\circ}$ C with end-labelled (TTAGGG)₃ or (TTGGGG)₄ respectively. Filters were washed in 3 $\times$ SSC at 42 $^{\circ}$ C for the (TTAGGG)₃ probe or 4 $\times$ SSC at 50 $^{\circ}$ C for the (TTGGGG)₄ probe.

could still be significant. As each cell contains 92 telomeres and the distribution of telomere lengths is wide, a loss of 2 kb from the mean length may imply a large increase in the proportion of cells missing TTAGGG from at least one telomere. The loss of even one telomere could cause the permanent cell-cycle arrest and chromosomal instability characteristic of senescent fibroblasts²². Most significantly, chromosomal abnormalities increase dramatically in the terminal phase of fibroblast cultures²³ and about 90% of these are dicentric attached at their telomeres²⁴. Loss of critical telomere sequences in late-passage fibroblasts could explain these data. Whether loss of telomeric DNA *in vivo* might contribute to the elevated (but still relatively low)

frequency of chromosomal rearrangements, particularly dicentrics, observed in older humans^{25,26} is uncertain.

Tumour cells are also characterized by shortened telomeres^{27,33} and increased frequency of aneuploidy, including telomeric associations²⁸. If loss of telomeric DNA ultimately causes cell-cycle arrest in normal cells, the final steps in this process may be blocked in immortalized cells. Whereas normal cells with relatively long telomeres and a senescent phenotype may contain little or no telomerase activity, tumour cells with short telomeres may have significant telomerase activity. Telomerase may therefore be an effective target for anti-tumour drugs.

In most unicellular eukaryotes, telomere length is a dynamic variable with both expansion and contraction events occurring under different culture conditions^{5,7,8}. Complex regulation of telomere length is not unexpected in immortal, single-cell organisms which experience dramatic changes in their environment. However, it is possible that mortal, somatic cells of higher eukaryotes lack mechanisms for telomere elongation. Telomeric sequences may be added exclusively in the germ line, thus explaining the observation that sperm telomeres are longer than telomeres from normal somatic cells^{17,21,29}, and these telomeres may simply be lost during the life span of an individual.

There are a number of possible mechanisms for loss of telomeric DNA during ageing, including incomplete replication, degradation of termini (specific or nonspecific), and unequal recombination coupled to selection of cells with shorter telomeres³⁰. Two features of our data are relevant to this question. First, the decrease in mean telomere length is about 50 bp per mean population doubling (Fig. 2a) and, second, the distribution does not change substantially with growth state or cell arrest. These data are most easily explained by incomplete copying of the template strands at their 3' termini. But the absence of detailed information about the mode of replication or degree of recombination at telomeres means that none of these mechanisms can be ruled out. Further research is required to determine the mechanism of telomere shortening in human fibroblasts and its significance to cellular senescence. □

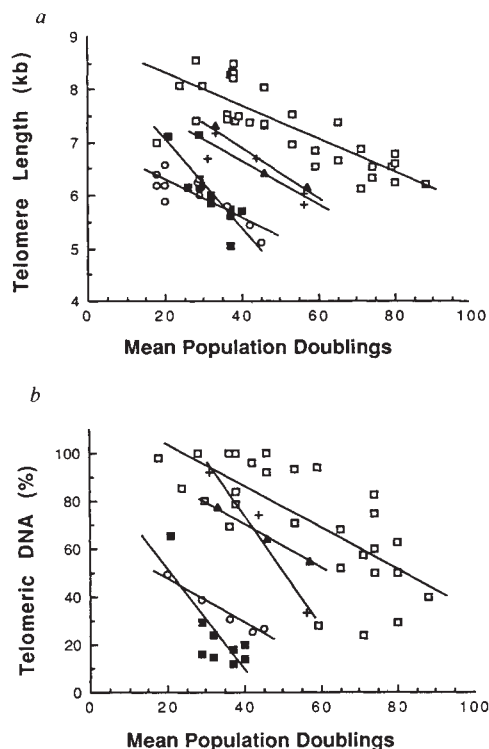


FIG. 2 Decrease in amount of telomeric DNA during replicative ageing of cultured human fibroblasts. Mean telomere length (a) and total telomeric DNA (b) are shown as a function of *in vitro* age. Symbols represent fetal cell strain HSC172 (□), newborn strain A30S (▲), young adult strain A38 (+), and old adult strains F001 (■) and F002 (○) (see Table 1). The slopes of the linear regression lines are: -31 bp MPD^{-1} , $-0.9\% \text{ MPD}^{-1}$ (HSC172); -48 bp MPD^{-1} , $-0.9\% \text{ MPD}^{-1}$ (A30S); -41 bp MPD^{-1} , $-2.3\% \text{ MPD}^{-1}$ (A38); -85 bp MPD^{-1} , $-2.1\% \text{ MPD}^{-1}$ (F001); and -37 bp MPD^{-1} , $-1.0\% \text{ MPD}^{-1}$ (F002). Analysis of HSC172 cells include data collected from two separate ampules of cells grown at different times.

METHODS. DNA digested with *MspI* and *RsaI* was quantified and $2 \mu\text{g}$ run in each lane of various 0.7% agarose gels for transfer to Nytran filters and hybridization with (TTAGGG)₃ as described in Fig. 1. Autoradiographs were generated without an intensifying screen using the linear response range of the film and scanned with a densitometer. The signals probably include hybridization to variant forms of TTAGGG. Output was digitized. Mean telomere length was defined as $\sum(\text{OD}_i)/\sum(\text{OD}_i/L_i)$, where OD_i is the densitometer output (arbitrary units) and L_i is the length of the DNA at position i . Sums were calculated over the range 3–17 kb. This calculation assumes that DNA transfers at equal efficiency from all points in the gel and that the number of target sequences (telomere repeats) per DNA fragment is proportional to DNA length. Although neither of these assumptions is strictly true, the data did not warrant a more complicated model. Qualitatively similar results were obtained with other assumptions and methods. To estimate loss of total telomeric DNA, the integrated signal ($\sum \text{OD}_i$ from 3–17 kb) was determined for each lane and, where possible, normalized to the signal from other Southern blots using a control probe. Lanes in which amounts other than $2 \mu\text{g}$ of DNA had been loaded were omitted from the analysis. Integrated signals for each autoradiograph were then expressed as a percentage of the signal from early passage HSC172 DNA on each autoradiograph as a standard.

Received 5 October, 1989; accepted 4 April 1990.

- Olovnikov, A. M. *J. theor. Biol.* **41**, 181–190 (1973).
- Watson, J. D. *Nature new Biol.* **239**, 197–201 (1972).
- Mueller, H. J. *Collecting Net* **13**, 181–198 (1938).
- McClintock, B. *Genetics* **26**, 234–282 (1941).
- Blackburn, E. H. & Szostak, J. W. *A. Rev. Biochem.* **53**, 163–194 (1984).
- Bernards, A., Michels, P. A. M., Lincke, C. R. & Borst, P. *Nature* **303**, 592–597 (1983).
- Shampay, J., Szostak, J. W. & Blackburn, E. H. *Nature* **310**, 154–157 (1984).
- Larson, D. D., Spangler, E. A. & Blackburn, E. H. *Cell* **50**, 477–483 (1987).
- Shampay, J. & Blackburn, E. H. *Proc. natn. Acad. Sci. U.S.A.* **85**, 534–538 (1988).
- Murray, A. W., Claus, T. E. & Szostak, J. W. *Molec. cell. Biol.* **8**, 4642–4650 (1988).
- Greider, C. W. & Blackburn, E. H. *Nature* **337**, 331–337 (1989).
- Greider, C. W. & Blackburn, E. H. *Cell* **43**, 405–413 (1985).
- Zahler, A. M. & Prescott, D. M. *Nucleic Acids Res.* **16**, 6953–6972 (1988).
- Shippen-Lentz, D. & Blackburn, E. H. *Molec. cell. Biol.* **9**, 2761–2764 (1989).
- Morin, G. B. *Cell* **59**, 521–529 (1989).
- Cooke, H. J. & Smith, B. A. *Cold Spring Harb. Symp. quant. Biol.* **51**, 213–219 (1986).
- Lundblad, V. & Szostak, J. W. *Cell* **57**, 633–643 (1989).
- Hayflick, L. & Moorhead, P. S. *Exp. Cell Res.* **25**, 585–621 (1961).
- Stanulis-Praeger, B. M. *Mech. Ag. Dev.* **38**, 1–48 (1987).
- Prashad, N. & Cutler, R. G. *Biochim. biophys. Acta* **418**, 1–23 (1976).
- Allshire, R. C., Dempster, M. & Hastie, N. D. *Nucleic Acids Res.* **17**, 4611–4627 (1989).
- Saksela, E. & Moorhead, P. S. *Proc. natn. Acad. Sci. U.S.A.* **50**, 390–395 (1963).
- Sherwood, S. W., Rush, D., Ellsworth, J. L. & Schimke, R. T. *Proc. natn. Acad. Sci. U.S.A.* **85**, 9086–9090 (1989).
- Benn, P. A. *Am. J. hum. Genet.* **28**, 465–473 (1976).
- Prieur, M. *et al. Hum. Genet.* **79**, 147–150 (1988).
- Bender, M. A., Preston, R. J., Leonard, R. C., Pyatt, B. E. & Gooch, P. C. *Mutat. Res.* **212**, 149–154 (1989).
- Hastie, N. D. & Allshire, R. C. *Trends Genet.* **5**, 326–331 (1989).
- Pathak, S., Wang, Z., Dhalway, M. K. & Sacks, P. C. *Cytogenet. Cell Genet.* **47**, 227–229 (1988).
- Cross, S. H., Allshire, R. C., McKay, S., McGill, N. & Cooke, H. J. *Nature* **338**, 771–774 (1989).
- Harley, C. B., Shmookler Reis, R. J. & Goldstein, S. J. *J. theor. Biol.* **94**, 1–12 (1982).
- Harley, C. B. *Meth. molec. Biol.* **5**, 25–32 (1990).
- Morgan, A. R., Lee, J. S., Pulleyblank, D. E., Murray, N. L. & Evans, D. H. *Nucleic Acids Res.* **7**, 547–569 (1979).
- de Lange, T. *et al. Molec. cell. Biol.* **10**, 518–527 (1990).

ACKNOWLEDGEMENTS. We thank R. Allshire and S. Goldstein for helpful discussions, C. Patterson for skin biopsies for strains F001 and F002, J. Regoezi for technical assistance, M. Buchwald for strain HSC172 and S. Goldstein and E. Moerman for DNA from strains A30S, A35 and A38. C.W.G. and A.B.F. thank Cold Spring Harbor for support. C.B.H. thanks the Medical Research Council of Canada.